UNIVERSITY OF LINCOLN

MASTERS THESIS

Toxic Behaviour in Multiplayer Competitive Video Games

Author:
Sam FARNDALE

Supervisor:
Dr. Chris HEADLEAND

A thesis submitted in fulfilment of the requirements
for the degree of Master of Science
in the

interAction Lab
School of Computer Science

January 2019

UNIVERSITY OF LINCOLN

Abstract

School of Computer Science

Master of Science

Toxic Behaviour in Multiplayer Competitive Video Games

by Sam FARNDALE

Toxic behaviours in online competitive video games are often discussed but very nebulously defined. This is despite toxic behaviours being rampant in competitive online video games with most matches seemingly containing some unacceptable behaviours. A focus is given to understanding feelings towards toxic behaviours from the perspective of gaming community members. Seeing these behaviours from a community perspective is vital to ensure that games create and enforce rules with their communities best intentions at heart. These may or may not align with video game rule sets provided by developers and publishers. A study was conducted first to inform on community experiences with toxic behaviours to increase understanding. This increased understanding and examples were used to create a further study to directly assess community members feelings on the behaviours given as examples. An attempt to categorise these behaviours was made using game rule sets and the accuracy of this assessed by community response. From these studies, conclusions can be drawn relating to real world experiences with toxic behaviour. It is found that toxic behaviours are heavily linked to the type of game being played with competition playing a key role. Several different categories of toxic behaviour are defined and assessed for community understanding and expectations. Suggestions are made to improve understanding of toxic behaviours in future studies and possible other avenues for future research.

Acknowledgements

Thank you to Dr Chris Headleand for being a fair and supportive project supervisor. You have continually ensured I maintain high standards throughout and taught me how to conduct this research to the best of my ability.

Thank you to my parents Tracy & Peter Farndale for their continued faith in me and my abilities with unwavering support in every way possible.

Thank you to all of my family and friends for their continued support from my bachelors all the way through to the completion of this thesis. Extra thanks to Danielle Scullion for helping to keep me relaxed throughout the year with an open door and open arms and Beatrice Wilson for keeping me motivated in the final months of this body of work.

Thank you to the University of Lincoln ICT department, especially Elaine Brown, and Auto Trader UK, especially Rob Chisholm, for providing me a supportive and flexible workplaces during this long process.

Contents

Nomenclature

- **Overwatch** - A first person shooter title released by Blizzard Entertainment on 24th May 2016 on PC, PS4 and Xbox One.
- **League of Legends** - A free to play MOBA title released by Riot Games Inc. on 27th October 2009 on PC.
- **Tom Clancy's Rainbow Six Siege** - A tactical first person shooter title released by Ubisoft on 1st December 2016 for PC, PS4 and Xbox One.
- **DLC (Downloadable Content)** - Any additional content added to a game post-release. This can be free or at a further cost.
- **MOBA (Multiplayer Online Battle Arena)** - A genre of competitive online video game where two teams compete to siege the others base with the eventual goal of destroying it to win.
- **RTS (Real-Time Strategy)** - A genre of competitive online video games where player/s compete against a number of opponents, controlling multiple units strategically with the goal of defeating their opponents.
- **FPS (First Person Shooter)** - A genre where players are placed in a first person camera view of a character and provided weaponry to shoot at opponents.
- **MMO (Massively Multiplayer Online)** - A genre consisting of very large playerbases all coexisting within the same game world/universe.
- **esports (Electronic Sports)** - Professional level competitive video games played in large events akin to a major sporting event.
- **Trolling** - Behaving in a playful manner in a game whilst not necessarily being toxic.

- **PvE (Player versus Environment)** - Games where players do not compete directly against one another. Players compete against environmental elements and non-player characters.
- **PvP (Player versus Player)** - Games where players directly compete with one another directly.
- **Personal Toxic Behaviour** - Toxic behaviours which are a personal attack on an individual or their beliefs.
- **Gameplay Related Toxic Behaviour** - Toxic behaviours which affect the gameplay experience of an online competitive video game for the participants.
- **Non-Player Character (NPC)** - A character in a video game that is not being controlled by a player.

Chapter 1

Introduction

1.1 Background

Anti-social and disinhibited behaviour in competitive online multiplayer video games has become a major problem for game developers and publishers in recent years. Developers have stated that development time is being actively taken away from other aspects of games to deal with toxic behaviours [67]. Attempts are being made to curb the perceived growth of toxic behaviours within games in an attempt to reduce their prevalence. The lead developer of *Overwatch* stated in September 2017, in the time taken for the development team to add reporting players to game console versions they could have made a match history system or game replay system. He also revealed that, 16 months after release, they had taken disciplinary action on over 480,000 accounts. In-game behaviour reporting systems like those developed for Overwatch allow players to report others in their games for breaches of a games behavioural guidelines. These reports can lead to various punishments for misbehaving players such as temporary bans, communication mutes or even permanent bans for serious transgressions.

Most online competitive multiplayer video games have an established set of rules which govern behaviour in a given match [30]. These behavioural guidelines may vary by developer, publisher, genre and individual game. The mechanics of a game may necessitates the inclusion of certain behavioural restrictions in some games but not others. For example in MOBA (Multiplayer Online

Battle Arena) games, leaving a match in progress is usually considered as unacceptable. This is not the case for many other online multiplayer video games such as *Starcraft 2* from the RTS (real-time strategy) genre or other games where team sizes are not large and game length is as long as a player chooses. Behaviour guidelines are put in place to try and ensure players respect the rules and behave in a civil manner during matches. Players are usually given the option to report players who do not follow these guidelines during or after matches. These in-game report systems are expected to be utilised by players in combination with automated systems to identify deviant behaviours.

Some games which are infamous for toxic communities being uncontrolled, such as *Tom Clancy's Rainbow Six Siege*, have recently been forced to take action. In this game, two teams of five are tasked with taking turns attacking and defending an objective until one team wins three rounds. The competitive environment created by this match type lead to serious incidents of toxic behaviour. These included frequent killing of team members and use of racial or homophobic slurs in the in-game chat. The game has a pre-existing system which catches and punishes cheaters but any other behaviours were unmonitored. This lead to Ubisoft announcing a drastic change with a new system added to the game to deal with toxic behaviours [33]. This new system instantly bans any user caught using homophobic and racial slurs or hate speech in matches. These bans scale up for each offence from 2 days to permanent. This permanent ban is a major change as this game is not free to play and so a player can be permanently banned from a game they previously purchased.

Many of the most popular online multiplayer video games on PC of recent years are based on a free to play model. As of February 2018, 4 of the top 10 most played PC games in the United States and Europe are free to play [50]. Free to play allow players access to much of the game and its content for free. Income is usually made by the sale of cosmetic items or characters within the game, depending on genre. These additional items may be considered DLC (downloadable content), additional extras which add new content or functionality to the base game which is entirely optional and often charged at a set price

separately. Games being available for free combined with low system requirements makes these games very accessible to most players. In free to play games, a player can have as many accounts as they like without having to repurchase the game so if an account is banned the player can simply make a new one. Doing this, however, does usually result in losing both earned content and any DLC purchases associated with the banned account.

Competitive online multiplayer titles almost always feature a specific competitive mode where each player is ranked on their performance. Usually a player can increase this rank by winning matches in competitive mode and lose rank by losing matches in competitive mode, tracked on some implementation of a points system. Ranked systems usually allow either an individual to play alone or for them to play with one or more friends as part of the same team. This competitive environment has been observed to increase anti-social behaviours when compared to casual game modes [65]. Versus casual games, individuals feel more invested in their rank and care about their progress at climbing rank and improving at the game. Usually in casual games individuals will play with groups of friends but this may not be possible due to limitations placed on competitive queues. Alternatively, many players choose to play these competitive games solo instead of with friends which can further change the way they treat the experience [69].

This competitive mode can often be taken even further with many of these games having active esports scenes for professional level play. Games such as *Overwatch*, *League of Legends* and *Dota 2* have official yearly tournaments organised by the game developer. At these events, esports personalities compete for vast prize pools in teams against other professional teams. These players often maintain this as a career option and full time, salaried profession. Players are often also paid portions of any prize money from tournaments [52]. Some colleges now also offer courses in competitive video games which can be chosen amongst other subjects or as a degree in itself [6]. Many players aim to improve their play in their favourite game to a point of becoming a professional and making a career out of their recreation.

1.2 Motivation

It is clear from the prevalence of toxic behaviours and increasing time demands to deal with it that research is needed in this area, both in understanding of the problem and development of possible solutions. This research is also vital with playing video games becoming both recreation for some and a profession for others. In March 2018 many companies with an interest in this formed a new organisation, the Fair Play Alliance [2]. This organisation is a collaboration of over 30 companies with an interest in curbing toxic behaviours through collaboration. The developers of *League of Legends*, *Overwatch* are members of this organisation along with games related companies such as Twitch, the largest game streaming website, and Discord Inc., creators of a popular communication platform for gamers. A new atmosphere of collaboration instead of competition further emphasises the size of the problem and the need for working together to understand and reduce toxic behaviours. This group also ties together the idea that being fair in a competitive video game is not simply playing fairly, but also treating others in your match fairly [2].

Games like *Tom Clancy's Rainbow Six Siege* recently being forced to act on toxic behaviour so long after release demonstrates that toxic behaviour cannot be ignored. Other games such as *Overwatch* have acknowledged having to change development priorities to accommodate for toxic behaviour within the playerbase. This even extends to older titles or games without behaviour monitoring which are now having it retroactively introduced. The necessity of this shows just how prominent the issues posed by toxic behaviour are to companies maintaining these competitive titles. Blizzard have stated that due to these behaviours they are having to take development time from other aspects of their titles [21]. A better understanding of toxic behaviours themselves will hopefully lead to less development time required to deal with them. This will allow for development focus to be placed more on the game itself instead of community behaviour within it and lead to more complete, polished titles and result in better games for everyone.

1.3 Research Questions

In an attempt to deliver further understanding of toxic behaviour, these research questions have been established:

1. What would be an accurate definition of toxic behaviour, if one is possible?

The most basic requirement of this research is to gather a greater understanding of toxic behaviour as a concept. Toxic behaviour is a very nebulous term with a great number of varying interpretations and a vast array of definitions. A firmer definition of toxic behaviour itself would prove useful for all involved.

2. How does a games community feel about toxic behaviour?

To understand why toxic behaviour is a problem, it must be understood why the community find it to be an issue. Toxic behaviour may be caused by in-game mechanics, human psychology, manipulation or a mix of these and other factors. A better understanding of these motivations and community feelings will help to increase understanding of toxic behaviour.

3. Are there different levels of severity in toxic behaviour?

Serious amounts of development time are being invested into curbing all toxic behaviours. A better understanding of prevention methods, both existing, past and theoretical, could aid in this process. A ranking of severities of different kinds of toxic behaviour could indicate what areas should be a primary focus or indicate a new approach could be necessary. This could hopefully result in reduced time and manpower requirements to deal with toxic behaviour.

1.4 Aim

The aim of this research is:

- Form a well rounded understanding of feelings towards toxic behaviour in competitive online video games to aid in reducing the negative impacts on online competitive video games communities.

1.5 Objectives

To achieve this aim, a serious of objective were defined:

- Understand the systems currently in place to deter toxicity in competitive online video games.
- Understand previous studies into the effects of playing competitive online video games.
- Find or derive definitions of toxic behaviour from a wide array of sources.
- Discover what form/s toxic behaviour can take.
- Discover whether there are different severities of toxic behaviour.

1.6 Chapter Overview

This body of work will be split into several chapters. These are as follows:

1.6.1 Chapter 2 - Literature

In this chapter, existing literature in the field is reviewed to grant a greater understanding of the subject area. Past research provides insight into interesting areas for future study or allows for gaps in understanding to be identified. Existing online competitive video game behavioural guidelines are also discussed. The systems used to enforce these existing behavioural guidelines are reviewed and possible issues identified.

1.6.2 Chapter 3 - Study 1

The first study conducted was an internet forum scrape for specific keywords in specific forums related to toxic behaviours in competitive online video games. This chapter details the study design and the reasons this approach was taken along with the methodology for data gathering. Results from the forum scrape are analysed using a thematic analysis and a linguistic analysis. Insights from these analyses are summarised and discussed.

1.6.3 Chapter 4 - Study 2

The second study was conducted as a custom web survey using data gathered from study one to create examples and generate further insight. Details are given on the basic survey design and the reasoning for design decisions made along with data analysis methodologies. Results were analysed for statistical significance and links drawn to confirm or deny hypothesis from study one.

1.6.4 Chapter 5 - Conclusions

Insights from all chapters are summarised into key points of intrigue or major significance. A comparison is made between actual findings and those initially expected results before studies were carried out. Areas where further understanding is necessary or further research is encouraged are given.

1.7 Contribution Statement

This research will ideally lead to a more solid definition of toxic behaviour in competitive online video games. The improved definition may be used in future studies to begin with a more solid understanding of the subject area. As this definition was derived from community members it should also prove useful for understanding gaming communities more completely. A greater understanding will also be given of the feelings players have in relation to toxic behaviours. Analysis of data directly related to different categories or forms of toxic behaviour should also allow for further studies to begin with these different variants more clearly understood.

Chapter 2

Literature

Existing research in this area often focuses on specific aspects of toxic behaviours or their effects. This may be enjoyment, anonymity, social interactions or aggression and relate to other similar methods of communication like social media. There is also a great amount of conflict between both the interpretations and conclusions found in various studies. Conflicting conclusions have also been generated as to whether video games can cause behavioural changes or even psychological harm to those playing them. There are also many meta reviews of other researchers findings further building on well rounded past studies [5].

2.1 Enjoyment

Players of different video game genres often have different criteria for enjoyment of their experience and satisfaction of progress made [36]. It has been shown in previous studies [44, 47, 11] that video games can have both positive and negative effects on a person's well-being depending on the individual players experiences. Consideration was given to problematic playing and the importance of separating this from the average player due to the differences that this can cause to reflect in the players experiences [39, 13]. Johnson's online questionnaire was used to determine where to focus on acquiring further data for analysis. The number of genres was narrowed to seven: Action Adventure, Action Role-Playing Game, Massively Multiplayer Online Role-Playing Game, Real-Time Strategy, Role Playing Game, First-Person Shooter and Multiplayer Online Battle Arena. These genres of video game were then compared

and contrasted over several categories such as frustration (how annoyed the game made the player at times), challenge (the perceived difficulty of the game) and immersion (how much a part of the games world the player feels).

Analysis of this data showed that Action Adventure games, Role Playing games and Action Role Playing games prove to be the most immersive within the study. This aligns with previous findings [35]. There is a clear link found between multiplayer games versus single player games on immersion. Single player experiences are rated as far more immersive than multiplayer experiences. It is thought this is because multiplayer games are more about social interaction and connecting with other people [14] as opposed to single player games focus on connecting with the in-game world. It was also found that games in the MOBA genre are an outlier showing different results from any other popular genres determined by the study. The MOBA genre was also shown to be more consistent in gameplay and mechanics between different titles within it. Several other factors were found to be of relevance such as the idea that as a more recently emerging genre, going through a proliferation in recent years, the MOBA style control scheme has not yet become second nature to many players. Other genres of game are more well established and share far more in common within their control schemes when compared to a MOBA. The measures used for player experience in Johnson's experiment show very little appeal for the genre which contrasts with their popularity [64]. It was concluded that the measures used do not adequately capture the components which draw a player to MOBAs.

The article then moves on to analysing MOBAs as the most popular genre of online video games in the world and a brief history of the genres creation as a mod. It was noted that very little research, at time of publication, had been done in relation to player experience in MOBAs. A comparison was done between other studies and journals to determine what factors may have been missed. There is a heavy focus on teamwork, communication and general social interaction provided by such games [44]. Past findings inspired a second study undertaken to interview specifically MOBA players to determine their motivations and experiences. They determined that competition is a valued

and important aspect of the players experience. This fits with the extremely large esports scene which has developed around MOBA titles over the past few years. It was also found that players gain satisfaction from noticeable improvements in their skill level and mastering mechanics presented in gameplay [47]. Teamwork was found to be another key motivator which keeps players enjoying their experience. This can take the shape of playing with friends for fun or generally performing well as a team or a mix of these and more possible factors [54]. Toxic behaviours in video games can prevent effective teamwork which would, following this analysis, decrease a players enjoyment of the game to a measurable extent [41]

2.2 Aggression & Competition

Ties have frequently been established between competitive video games and violent or aggressive tendencies [1]. Adachi attempts to link the actions a player performs within a game to the general competitive nature of a game (or lack thereof) and certain violent characteristics expressed by players. This was determined by testing for all factors independently from one another to avoid any cross-over which would invalidate gathered data for a certain variable within the testing. It is stated that many other tests in the past have had varying numbers of competitive objectives or not directly compared violent and non-violent games [72]. It is also stated that many former studies of violent video games have used very unrealistic and animated violence. The effects this could have on the final results of the study have never been measured. Adachi also criticises commonly used methods for determining levels of aggression in relation to video games. Adachi used the Taylor Competitive Reaction Time Test (TCRTT) as part of this study. TCRTT is often used in laboratory studies into aggression by asking participants to fire sound blasts at fictional opponents. The use of the Taylor Competitive Reaction Time Test (TCRTT) is heavily criticised for the ambiguous nature of the participants motivations and a proven lack of validity [42]. The ambiguousness is argued to be a key preventative factor as the difference between competitiveness and aggression cannot be ascertained. The

validity is questioned as past tests have shown there to be little link between results of the TCRTT and behavioural histories of participants related to violence and aggression [23]. They instead find the Hot Sauce Paradigm to be a more reliable testing methodology. This method involves asking participants to give a portion of hot sauce to a person they know hates hot sauce. The person will have to eat all of it and the size of the portion is determined by the participant.

Several past experiments using games such as *Halo* and *Super Mario Bros.* [4, 63] have attempted to discover a link between how competitive an objective is and how aggressive players appear after playing a session. This level of aggression is determined using various methodologies such as word association tasks. The biggest difference was measured between those completing competitive tasks and those completing cooperative tasks. To attempt to measure the effect of violence, Anderson uses five games which were determined to be of both comparable and different levels of violence and competitiveness between them. These games were *Conan*, *Fuel*, *Mortal Kombat versus DC Universe*, *Marble Blast Ultra* and *Left 4 Dead 2*. The experiment conducted here, much like previous referenced experiments, shows that competitive objectives have a significant and measurable physical effect on the player. These physical affects measured include heart rate and perspiration. It is also shown that playing a game which is both competitive and violent further amplifies the measured physical effects on the player. This demonstrates that the level of violence in a game causes further changes in the physical characteristics of a person in this experiment [1].

The results shows real, tangible physical repercussions to being competitively violent within a game compared to a lesser increase for being competitively non-violent. The impact of a competitive, non violent game is still measurable but violence is clearly a factor which can influence the players response. A clear correlation is found here between the level of violence in a competitive game and the physical reaction shown by the participants body. The conclusion drawn is that it is not violence, but competitiveness which leads to a short-term increase in violent behaviours.

Other studies have also found increased levels of aggression when players

are competing versus cooperating in online video games [18]. Eastin makes a point of considering the depth of modern video games and advanced concepts including group dynamics and gameplay communication. Past studies often do not include these as they were not commonly present in games at the time [3]. This is in an effort to generate a more accurate representation of group play in modern games than past studies which may have oversimplified the variables involved [71]. This study found that behaviour was similar in small groups but when moving from a two-person or four-person to a six-person group the level of hostility increased by a greater amount. A hypothesis suggested to explain this is that the number of relationships involved in a six-person group versus a smaller group increases multiplicatively. This increases the perceived anonymity of the participants.

This increased anonymity leads to increased levels of predicted hostility [73]. Many online games feature game modes with six or more participants which would lead to greatly increased levels of perceived anonymity and hostility, following this hypothesis. Eastin also found little to no impact on state hostility from in-game violence such as killing. It is suggested here that, again, as games become more sophisticated researchers need to move on from very simplistic interpretations of violent gameplay in video games. Modern games are much more complicated entities with many modern systems and facets which need to be accounted for. This concurs with the findings from other studies [1] that violence in itself is not as simple of a variable as it has been considered in may past studies. It is also possible that a multiplayer environments dilute some previously found aggressive tenancies and short term anger when compared to single player experiences [23].

Studies have also been conducted on the effects of profanity on player aggression [34]. Profanity can be considered a form of verbal aggression [37] which the player experiences when playing approximately 1 in 5 video games at the time of Ivory's study. It is hypothesised that the effects of profanity may include increased expectations of hostility and more violent thoughts by the players. Ivory found that the use of profanity by both video game protagonists

and antagonists increased a players expectations of hostility. There was, however, no perceivable difference between a protagonist or an antagonist character using profanity. The study clearly demonstrates that exposing a player to profanity causes a change in players expectations of their video game experience. This could be a further factor towards past findings of video game driven aggression which has not been considered in past [55]. It was also found that profanity from antagonists nor protagonists appeared to have little effect on aggressive player thoughts. It is also noted that it is hard to isolate the effects of profanity specifically. Every game tested which used profanity also featured violence and would be considered a violent game. This creates great difficulty in attempting to measure the effects of profanity alone. Variation could also be found between genres or between NPC vs players use of profanity in further studies.

2.3 Psychological Effects of Video Games

The frequent playing of video games has been linked to serious behavioural issues in people such as depression and aggression [11]. The data used for the Brunborg study was provided by a large, census style survey given to various ages of adolescents in Norway. The survey was intended to be administered to the same students once and repeated two years later in an attempt to measure the change over time as the student matured. "Young in Norway 2010" and "Young in Norway 2012" provide an extremely large amount of data as 89 schools in Norway contributed. There were 8,356 participants in 2010 and 2,450 of these same participants further participated in 2012. These responses are used to generate measures from several different questions to attempt to create metrics to base comparisons upon. The sample size of this experiment is very large compared to many comparable studies previously [72, 4].

Brunborg found a significant negative correlation between amount of time spent playing video games and depression, serious and aggressive behaviour problems. This changes drastically when the level of video game time rises to addiction where the relationships are almost entirely reversed. The trend

shows that playing some video games is positive but playing excessive video games has the opposite, negative, effect on mental health. There is a positive correlation between video game addiction and serious, covert and aggression behavioural problems. A heavy engagement with video games but no addiction showed no discernible negative effects on an individual's mental health or relation to behavioural issues. The authors heavily encourage that future studies continue to acknowledge the separation between playing large amounts of video games and addiction due to stark contrasts found between the effects of each. These results concur with past studies [58] but contrast with others [29, 28] who found that some negative mental health effects could be found in early adolescents.

Meta reviews [5] have concluded that games do have a psychological and emotional effect on the player. Some studies have even found that factors from video games may cause a player to dehumanise their opponent [7]. Anderson calls into question the conclusions of more recent studies which appear to contradict this and claim that links to aggression have been overestimated [22]. This paper was an attempt at creating a detailed and unbiased meta-analysis in response to the perceived bias in other attempts. Effort was also made to account for cultural differences and the effects this may have on aggression. Several variables were established such as categories of behaviour (pro-social behaviour, physiological arousal etc.), culture (east or west) and the research design (Cross-sectional, experimental etc.). These were used to quantify aspects of the large number of other studies used to conduct this meta-analysis.

The analysis found a significant correlation between exposure to violent video games and higher levels of aggressive behaviour. It was also found that playing violent video games has a longitudinal effect, showing that it can lead to increased aggression over time. Results also imply that children may be more susceptible to this increase in aggression however it was noted that more studies into the matter are needed to confirm this. Anderson claims that the public debate should move on from finding a link to mitigating its effects. He also believes that more focus should be given to pro-social games as they have been linked to positive behaviours in much the same way as violent video games and

negative behaviours [27].

2.4 Toxic Behaviour

2.4.1 Enforcement of Behavioural Guidelines

A key feature of many toxic behaviours, as defined by behavioural guidelines for games, is that they are performed intentionally [19, 30]. It is also stated in most behavioural guidelines that being bad at the game is not a reportable offence. There is often some confusing within communities as to whether a bad player can be reported for breaking behavioural guidelines. It can sometimes be very difficult to determine whether poor performance is intentional or unintentional from observing other members of your team or an enemy team play the game. However, these behaviours are often accompanied by toxic behaviours exposed through in-game communication methods which make detection easier [40].

There have been several public instances of high profile personalities or players from games receiving punishments [59]. Some *League of Legends* personalities have been publicly banned as an individual and not an account. This means that if the developer ever discovers the individual has an account, they will take action to ban it immediately [61]. There have also been several instances of *Overwatch* players receiving punishments during live streams [56]. These punishments for toxic behaviour are usually bans for a given period of days and deny the player access until the time is up. This denies the player from playing or streaming the game for the duration as punishment. *Overwatch* esports players within the *Overwatch League* are held to strict behavioural rules at all times including outside of official events [25]. In *Tom Clancy's Rainbow Six Siege* a player is instantly banned for toxic behaviour in the chat and removed from a game in progress instantly without being reported [17]. This means that other players in the match are sometimes left a player down due to the toxic individual being removed immediately. If a replacement is not found, the team may lose the game purely due to being a member down on their team.

2.4.2 Why Does Toxic Behaviour Occur

Other areas of literature place a key focus on toxic behaviour itself [41]. Kwak focuses this in their paper on bullying and toxic behaviours in team competitive online games. The data used in this study is from a legacy League of Legends system called Tribunal [8]. This system allowed high levelled players of the game to review reports on other players. Players could then decide to punish or pardon the player. The final decision was made on consensus of Tribunal users and users who were part of this consensus received a small, in game reward. This system was taken down for maintenance by Riot Games in early 2014 and has remained offline since. Riot Games have stated this is due to the system being slow and inefficient as well as sometimes being very inaccurate due to participation rewards for players [26]. It has since been replaced by automated systems which reviews player reports and acts automatically on some behaviours such as hate speech and threats of harm. [43]. Data from this system over three League of Legends server regions (North America, Europe West and Korea) was used as a data set for analysis of player interactions. A strong focus was given to the vague nature of toxic behaviour, bystander effect [12] and the differences between playing with or against another player and report frequency. The bystander effect states that an individual is less likely to offer help to someone in need when others are present. Other factors such as a players gender may also have an effect on toxic behaviours [57].

The large data set used allowed for many statistics to be generated in relation to proposed hypotheses. It was found that if allies ask for a member of their team to be reported, that individual was 16 times more likely to receive a report from opponents. This is interesting as, since this report was published, Riot Games has confirmed that asking opponents to report an ally is considered a negative behaviour in itself [31]. This seems to show that the bystander effect is negated by a request from the toxic players allies. Kwak also found that more reports were made for more concrete toxic behaviours, such as intentional feeding and assisting enemy team, when compared to more vaguely defined toxic behaviours such as hate speech or offensive language. The ratio of pardon votes

on the Tribunal, particularly on the Korean server, shows that there was a clear difference between what is considered toxic between players. This appears to confirm the vague definition of many toxic behaviours.

2.4.3 Identifying Toxicity

The language used in online competitive video game matches can provide insight into the mindset of a player. This language may also show toxic behaviours from players in the match which can be analysed from a linguistic standpoint [40]. The data set used from this study is from the *League of Legends* Tribunal containing 590,000 cases. The same data set was been used for other studies conducted around this period due to its size and level of detail [41]. It is established early that toxic behaviour is bad behaviour which violates social norms and causes harm. A great deal of emphasis is placed on the idea that toxic behaviour causes harm far longer than the duration of a match. It is noted that offensive language and verbal abuse have been shown to be the most common toxic behaviours in previous work [40].

Kwak's aim is to categorise toxic behaviours linguistically and then try and discover how a player turns into a toxic player. Initially an analysis is conducted into the volume of chat messages over a match. It is found that there are three distinct levels of communication throughout a *League of Legends* match depending on the phase of the match. Players tend to communicate often in the early stages of the game when they are isolated on their lanes. The amount of communication lowers as the mid game begins and then raises again at the point of the endgame where teams plan out movements. It is found that at the start of a match the typical player speaks more than a toxic player. This changes after a short time and toxic players appear to chat more frequently than a typical player until the late stages of a game. This may be as toxic players gain more reason to speak up after gameplay has occurred which gives them something to complain about to the rest of their team.

The usage of words between a toxic and a typical player is also analysed. Some words are identified as exclusively used by typical players and explicitly used by toxic players. It is found that most verbal abuse from toxic players

takes place in the later stages of the game. This is presumed to be because in the later stages of the game it becomes more clear whether your team will win or lose the game. As it becomes clear that the game will likely be lost, a toxic player begins to be more abusive as this dawns on them. The distribution throughout the game also identifies the different natures associated with the different categories of toxic behaviour. In the early game toxic behaviour is often threats of intentionally feeding the enemy team from a toxic player. As the game moves to the mid and end game it is more common to observe verbal abuse and messages blaming the players from the bottom lane for losing the game. This is because in *League of Legends* the bottom lane typically contains the characters who are expected to be most powerful in the late game.

A suggestion is also made to have some form of system which would detect and warn a player when they are showing toxic behaviour. This would give a toxic player a chance to reform and change their behaviour during the game before they continue and are punished. It is also suggested that studies should be conducted into the flow of messages within a game instead of simply looking at language used. Kwak also notes that it is important to make the discrimination between toxic behaviour and 'trash talk'. This is because trash talk is considered an important part of immersion in the game and sometimes used by typical players. Any system which warns a toxic player of their behaviour must also be able to discern trash talk and be flexible enough to allow some of it to avoid breaking player immersion. Trash talk is done in a playful and friendly manner without any intent to have a harmful impact on other players which is a key difference between trash talk and toxic behaviours.

2.4.4 In-Game Performance

Another study builds on many of these findings and attempts to assess the impacts on other players of *League of Legends* [49]. This study focuses on the conversation topics within games and how they relate to others performance considering what they refer to as "toxic contamination". The data used for this study was from the same dataset as the Kwak study detailed previously but only records from the North American server were used. This data was used

to investigate toxic topics of conversation and their relation to toxic behaviours and toxic contamination. A performance metric for players was used based on kill, death and assist values included with the Tribunal data [8]. The metric being created in this way allows for its use to evaluate performance of individuals, whole teams or other subsets of players. The Tribunal data was analysed and then aggregated into subsets.

Neto concluded that the interactions between players in matches is very important and that topics discussed have a strong correlation to team performance. Teams that focus communication on teamwork and respect tend to have better performance in matches and lower levels of toxic contamination. Other teams with negative topics in conversation are shown to be linked with increased stress levels and decreased match performance for players. It was found that this toxic contamination can establish a vicious cycle that increases a players predisposition to toxic behaviour. This is very important as evidence of toxic behaviour propagating itself through toxic contamination could explain how rampant the problem with toxicity has become for many competitive multiplayer video games. The recommendation is made for games to have systems in place that recommend players take a break when negative communications are detected from them instead of immediate punishments. Comment was also made to several instances of players identified as toxic communicating in a positive way and attempts to find probable cause could be made.

2.4.5 Types of Toxic Behaviour

An attempt to determine the different types of toxic behaviour that exist is made by Rubin and Camm [62]. Rubin refers to these behaviours using another term for toxic behaviour, griefing behaviour. The aim of this study is to understand the different types of griefing in an empirical way to raise awareness of these behaviours. The study was conducted as an analysis of an internet forum followed by structured interviews. These were used to determine how players define griefing, players attitudes to griefing and whether or not they feel these behaviours are deceptive, It is determined that the categories determined previously by Foo in his past paper [24] are accurate for this studies findings with

one additional category. These are:

- Harassment
- Power imposition
- Scamming
- Greed play
- Casual griefing

Casual griefing was added to account for non-malicious acts of griefing. These behaviours are often referred to as trolling. Rubin further suggests adding another 6 dimensions to grade griefing behaviours:

- Intent to disrupt
- Maliciousness
- Lucrativeness
- Means
- Deceptiveness
- Emotional perception

It is also noted that there are several conflicts between study findings based on forum posts and those based on structured interviews. Interviewees did not often identify griefing as deception but forum posts were more focused on the deceptive aspects of griefing. This leads to a conclusion that griefing is a complex phenomenon and very open to different interpretations. These definitions are very thorough however not very relatable to the categories that players associate with toxic behaviours as defined by report systems in games. Players may find it hard to categorise the behaviours they experience and already categorise in a system they are familiar with using this entirely new system. There is also little granularity in categories such as harassment. Harassment can take many forms and could easily have subcategories.

2.5 Insights

It is made clear by several studies [47, 36, 54] into levels of player enjoyment from video games that the enjoyment is subjective to the individual. It also appears that different genres of video game provide enjoyment for different reasons and in different ways [36]. Some games are enjoyable because they are immersive and the player becomes part of the in-game world. Other games are enjoyable because of the level of complexity and challenge offered. Some games may be enjoyable to some purely because of the social interaction with others experienced through the act of playing. These social gamers may even find more enjoyment from the interaction than the games themselves. Genres of video game like MOBAs are found to offer enjoyment in a way which is not necessarily comparable to many other genres [47]. Additional studies have been conducted to discover why players enjoy games in the MOBA genre when, compared to other games, they do not fulfil factors established for player enjoyment. Video games clearly offer entertainment value in many different forms throughout all genres with varying levels of similarity between them.

There is a very large crossover in literature between extensive playing of video games and video game addiction [11]. Some papers make a significant effort to separate the two and to find the effects of each distinctly from one another. Papers which do this seem to find a drastic difference in the effects of playing large amounts of video games and addiction. They find that playing large amounts of video games is mostly beneficial to mental health, stress levels and overall happiness. However, addiction takes on negative effects more typically seen with other addictions. Individuals suffering from video game addiction appear to often be more stressed, depressed and allow video games to take over other life priorities. They also show more mental health issues such as depression over an individual who does not suffer from an addiction. This could mean that those addicted to video games are more likely to show negative behaviours in matches due to their negative frame of mind. More casual players would seem to be less susceptible to this as they do not have such negative mental issues as prevalently as those with addiction.

Violent video games are often perceived in media to increase levels of aggression in individuals who play them. Many academic studies, however, vary on this subject and find both this to be true and conflict with this view [51, 1]. Some studies have found that playing an aggressive video game increases predicted hostility of players. They have also found physiological evidence of this effect such as increased heart rates. Other studies have found, contrasting, that playing violent video games decreases the violent tendencies of an individual. Those tested have shown lower levels of anxiety and state hostility after playing a violent video game. This is believed to be because the violent video game provides an outlet for any violent tendencies in a safe environment. With this outlet, these tendencies are less pronounced in an individual during their everyday life.

Studies have also been conducted into the effects of profanity and offensive language on a players feelings in a video game [34, 66]. These studies find that, although hard to isolate in a testing environment, there is little change to aggressive cognition in a player. It is also found that the character who is using the profanity, be it protagonist or antagonist, does not show any measurable difference. However, it was shown that the use of profanity changes a players expectations of a video game. This could lead the player to treat a game with profanity in a different way than a game without profanity. It could also lead to a different mindset when approaching games with and without profanity which could affect aggressive cognition.

Many studies provide a focus on toxic behaviours and attempt to increase understanding of them. Some of these have found strong links between levels of toxic behaviour in a given match and team performance [54, 41]. It was found that a team which had noted toxic behaviours occurring within it performed noticeably and measurably worse than a team without. The toxic behaviour may dishearten the players on the team or cause them to become toxic themselves and initiate a downward spiral of performance. It is also indicated that toxic behaviours in a match increase the stress levels of the players taking part in the match. These stress levels may also be linked to team performance however this assertion is never tested. Many studies also note the fine line between toxic

behaviours and just being rude or demonstrating bad manners [40, 64]. Hate speech is widely accepted as toxic behaviour by the wider community however being rude to another player in a match may or may not be. These actions are more open to interpretation and so are sometimes classed as toxic behaviour but sometimes dismissed. A clearer definition is needed which helps identify for people who are unsure whether or not an example they have is toxic behaviour or not.

2.6 Industry Grand Challenges

2.6.1 Large Playerbases

At its peak, *League of Legends* had over 67 million active players each month with 7.5 million playing at once during peak hours [60]. This size of playerbase makes the typical moderation methods used in places such as online forums entirely infeasible. These standard moderation methods usually involve having a team of staff who manually review posts by users. A system like this for competitive multiplayer video games would involve a team of staff reviewing data about matches which take place. Each of the unique daily players likely takes part in more than one match on a given day. The sheer number of matches taking place daily would be insurmountable for a team of staff of a feasible size to review.

The same problem is present for many of the more popular online competitive titles. Games such as *Dota 2* and *PlayerUnknown's Battlegrounds* regularly reach or surpass 1 million unique players at peak a day [16]. Each game with a large playerbase would require an extremely large team of staff to review reports given by users and any automated flags. It is not feasible for each game company to hire such a large team for their game purely to review behaviours. This means that the traditional online moderation model is not applicable to moderation of behaviour in video game matches. As this is the case, alternative approaches which are more feasible must be considered.

2.6.2 Inconsistent Community-based Systems

Many alternative approaches have been attempted in the past to help with moderation of competitive video game matches. In the past *League of Legends* used a system called *The Tribunal* [43]. The data presented would show any toxic behaviours such as hate speech using in-game communication methods or gameplay related toxic behaviours such as intentional feeding. *The Tribunal* ran by majority rule, meaning that the most popular action between punish and pardon would be taken. This system allowed for the community to determine, following a set of guidelines, whether other community members behaviour was acceptable. Players who voted in the majority were rewarded with a small amount of in-game currency and the number of cases per day was limited to 10. *The Tribunal* was later shut down in favour of an automated system developed to punish toxic individuals [68].

Another alternative, similar system is still used in *Counter-Strike: Global Offensive* is their *Overwatch* system. This system allows players of an above average rank in the competitive matchmaking mode to review in-game replays of reported players. These replays are shown from the reported players perspective so they can be monitored during the match. At the end of the replay the observing player is asked to decide if the reported player was breaking game rules in any way. The observer votes whether or not a player has violated any rules in four categories of behaviour. These are: 'Aim Assistance', 'Vision Assistance', 'Other External Assistance' and 'Griefing'. As can be seen by these categories, the observer is monitoring for cheating and bad manners but cannot see instances of other toxic behaviours like hate speech.

These systems place the weight of making the final decision upon the wider community of the game. This requires they understand, comply and sometimes even agree with the rules in place. A player participating in this peer review system is free to choose differently if they do not personally agree with a rule of the game. They are also free to intentionally attempt to sabotage the system or attempt to quickly farm the system for any rewards given for its use. These key flaws have lead to the dissolution of systems like *The Tribunal* in favour of these

automated systems [26]. These automated systems are easier to manage and entirely under the control of the game developer or publisher who can choose and enforce rules as they see fit. They can also be adapted quickly to any rules changes that are made.

2.6.3 Inaccurate Automated Systems

The automated systems which are often used are, however, not infallible in themselves. An automated behaviour monitoring and punishment system was implemented into *Tom Clancy's Rainbow Six Siege* in early 2018 [33]. This system was put in place as the first step to lowering the levels of the toxicity present in the game which have previously been unmonitored. Several hurdles have been encountered by this system since its implementation [17]. The system has had a very high rate of false positives due to slang words in one region being considered toxic terminology in others. A notable example is English players of the game have experienced instant bans for referring to a cigarette as a fag. Although common slang for a cigarette in British English, fag is often considered a homophobic slur in American English. The automatic system has no understanding of context and so simply sees the use of the word and performs an instant temporary ban. Other terminology has also experienced these false positives.

Chapter 3

Study 1

3.1 Introduction

To facilitate gathering data for analysis, studies must be conducted. This study was aimed to gather well-rounded data on the domain of toxic behaviours in video games. As such, gathering data from areas where gamers who participate in these communities congregate seemed to be a logical choice. Communities built upon more engaged players of games where toxic behaviours may occur may contain vital data on understanding these behaviours from the perspective of the community at large. There may also be outliers in the community who demonstrate differences from the consensus and may prompt discussion. This basic data on preexisting communities provides the foundation which analyses, further studies and possible insights are be built upon.

3.2 Design

The first study was designed to gather community members experiences of toxic behaviour in their preferred online competitive video games. This data was to be gathered from online sources using some form of scrape to collect data en masse. The data selection was made based upon the post content. Posts which discuss toxic behaviours were taken by the scrape whilst others were ignored. A mixture of different forums for different games will allow for well rounded overview of toxic behaviours in general. Different games or genres of game may have very different gameplay related toxic behaviours. Individual

related toxic behaviours were expected to be generic and apply throughout all titles.

3.3 Methodology

The study was conducted on the website Reddit as the sole source of community posts and interactions. This is because Reddit has active communities for many popular video games [32]. Posts were gathered over a long period of time using a listener to scrape posts which contained certain key terms. The terms were: 'toxic', 'ban', 'report', 'troll', 'grief' and 'trash' The words 'ban' and 'report' were chosen as they are the terms used to refer to behaviour and punishment systems within games, specifically the parts which a general user may experience [19, 30]. 'Troll' and 'grief' were chosen as they are well established words used to refer to deviant behaviours within video games [38, 24]. 'Toxic' is a more recent development but now the accepted term for deviant behaviours against the rules of an online multiplayer game [41, 45]. 'Trash' was chosen during the testing stage of the script before the final scrape was performed. Including this word will help to create a well rounded insight into behaviours considered deviant by the community but which are not necessarily toxic by a games behavioural guidelines. It was chosen because many posts were observed including users perceiving toxic behaviours from bad players who they describe as 'trash' but who were not necessarily being toxic. These players were often shown to perhaps be bad or new at the game but others would consider their lack of performance as toxic, Most rule sets for games define being bad at the game as acceptable behaviour yet the boundary between this and certain toxic behaviours, such as assisting the enemy team, can be hard to define. Data on the use of this term may help to clarify this boundary and inform on community feelings around this topic.

These posts were stored in a spreadsheet and newer entries were appended to the top as they were chosen. This took place on several Reddit pages, or subreddits, for popular competitive multiplayer video games. These subreddits were: */r/leagueoflegends*, */r/Overwatch*, */r/Rainbow6*, */r/DotA2*, */r/PUBATTLEGROUNDS*,

/r/csgo, */r/Fortnite* and */r/Gaming*. These games were chosen as they are all active, popular competitive online multiplayer video games. These all have consistently large playerbases and online presence. Some of this is demonstrated in large followings on their forums and subreddit pages. These very active communities should be able to provide a large amount of posts which are also relevant with some containing insight. Over this long period of time it was expected that a large number of posts referring to toxic behaviour would be found. These posts would be spread between various popular titles selected to be monitored.

The inductive nature of this study will allow for generalised conclusions to be drawn. This study is formed by passive observation of gaming communities As the observations are passive, there is no influence on data gathered by the collection. Gathered observations will be without any guidance or influence from the study itself taking place. Observations are conducted of the post authors in their natural environment and without knowledge of observation. This should ensure that data gathered is accurate to posters true feelings and without any bias introduced. These inductive findings can lead to conclusions about the post authors feelings and opinions. Analysis of this data may lead to further conclusions formed from correlations or other apparent factors.

This analysis will take the form of a thematic analysis and a linguistic analysis. The thematic analysis will reveal common themes among the observed community posts. These references can be categorised into nodes and subnodes to identify themes and subthemes. These themes will be generic categories of references and behaviour. Subthemes will be more detailed and categorised examples within these themes [10]. This analysis will be performed using NVivo 11. The second form of analysis, linguistic analysis, will be used to determine the emotional state of posters. The inductive nature of this study allows for this to be observed in a natural environment to the poster. Linguistic analysis will reveal whether the poster feels positive or negative affect towards their experience. The analysis will also reveal whether the poster is using honest and open language or trying to deceive or hide information [40]. This will be performed using LIWC.

There are advantages and disadvantages to conducting an automated analysis [9]. The analysis is free of any bias which may be introduced by analysis conducted by a human. The process is also much more efficient at analysing large amounts of text. The LIWC system is already trained and so does not require additional preparation However, there are also some disadvantages. The analysis, much like behaviour monitoring systems, does not understand context or nuances such as sarcasm which may be present. The linguistic analysis may also miss any known hints towards deception if they are not outright stated. This could lead to some examples being misconstrued or missed.

3.4 Ethical Considerations

Scraping internet forums comes with many ethical factors that must be considered [20]. Before scraping, moderators were asked for the subreddits scraped if this research was okay to be performed. This was to ensure that the moderation team in charge of the subreddit were okay with the data they control being used in this fashion. Also, during the scrape many posts which were gathered were soon deleted by moderators of the subreddits. This was because they breached the community guidelines of the specific subreddit. Many do not allow posts where a user simply complains about receiving punishment for their behaviour. Due to the nature of the scrape these posts were still retained in the data set for analysis. As the initial posts were made by the author to be publicly accessible to anyone on the internet this should not be an issue. The removed posts will still be unavailable on the subreddit after being removed by the moderators so this should also not cause issue for them.

3.5 Implementation

The Reddit API was used to create a Python script which gathered posts related to some of the most popular online games. Posts featuring certain keywords were scraped from the website and collected in a spreadsheet. As stated previously, the scraped keywords were ‘toxic’, ‘ban’, ‘report’, ‘troll’, ‘grief’ and

'trash'. This script utilises the *PRAW* and *pygsheets* Python packages to allow for simple utilisation of the Reddit and Sheets APIs. Both of these APIs are fairly fully featured and allow for integration with site content after some security steps. Reddit and Sheets first require the script to be registered on their sites APIs respectively to allow access.

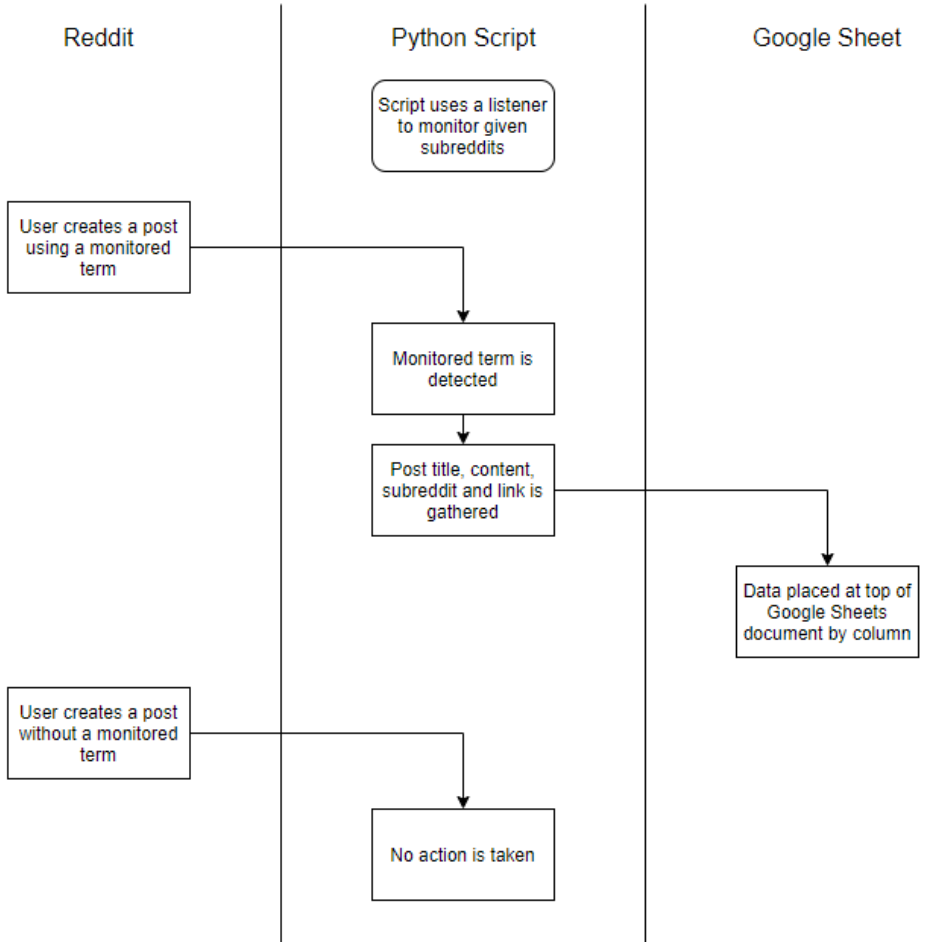

FIGURE 3.1: Reddit scrape data flow diagram

3.6 Results

2032 posts were collected over an 18 day period, from November 6th to November 24th 2017. This time period was used as the goal number of posts to analyse was 2000 or more due to many posts in initial observations having little to analyse by being too brief. They include chat logs, personal accounts, emotional statements and requests to the communities of selected games: *League of Legends*, *Overwatch*, *Rainbow Six: Siege*, *Dota 2*, *PlayerUnknown's Battlegrounds* and

Counter-Strike: Global Offensive. Other subreddits were included in the script parameters however no posts were scraped during the runtime of the script and, so, no posts were available from them: *Fortnite* and *Gaming*. This may be because Fortnite feature very little toxic behaviour, the community address it in a different fashion or some other unknown factor. This is interesting as Fortnite features both text and voice chat for its players. The Gaming subreddit is very generic so it is doubtful a player would complain about a specific game here.

Once the scrape was concluded the dataset was manually cleaned to remove any irrelevant data. The majority of the posts were extremely short or full of expletives whilst containing little substance or usable data. Posts left in the dataset include detailed reports of toxic experiences and relevant chat logs. Of the 2032 posts, 98 were deemed suitable for use in further analysis. These posts fit the following criteria:

- Posts must contain some data about the match involved.
- Posts need to be longer than a few sentences.
- Posts need to contain specific references to some perceived deviant behaviour within the game.

These suitable posts contain various players experiences in game matches with toxic players on their own or the enemy teams. The posts come in the form of user stories and chat logs. Some also include interactions with support agents from the game in question if the player was punished and challenged the judgement.

3.6.1 Thematic Analysis

Theme Nodes

A thematic analysis was conducted on these relevant posts by coding the titles and content of each post using NVivo 11. NVivo is a piece of software designed for conducting qualitative data analysis on text-based datasets. The posts were fed into the software, coded and then nodes were created to categorise them. Nodes were created to classify examples of relevant reference within the posts:

- **Reports & Punishment** - References made to reporting players or sanctions placed on players due to reports.
- **Toxic Behaviours** - Examples of any deviant or toxic behaviours.
- **Communications** - References to use of in-game communication methods.
- **Experience** - References to an individuals experience of playing the game.
- **Feelings** - References to any emotions experienced in relation to the game.
- **References to Developer** - Any discussion about the developer/publisher of the game.

These were chosen as they account for key topics both observed in posts and prevalent in relevant literature.

Theme Subnodes

Subnodes were then created within these to further differentiate between identified examples. Word frequency and text search queries were also used to discover any common themes between posts. Words were examined and compared using word trees and cluster analysis to determine context they are used in and word similarity.

These categories and the number of references within them are given below:

A significant number of different toxic behaviours were identified within the toxic behaviours category by the analysis process. Listed below is the type identified and the number of references made:

- **Assisting opponents** (104) - Hampering your own teams chances of victory by helping the enemy team or stopping your own team from achieving an objective. Commonly called 'throwing' in some team based games such as *Overwatch*.
- **In-Game suicide** (84) - Intentionally dying to assist the enemy team and help them win the game against you. Commonly known as 'inting' in some team based games such as *League of Legends*.

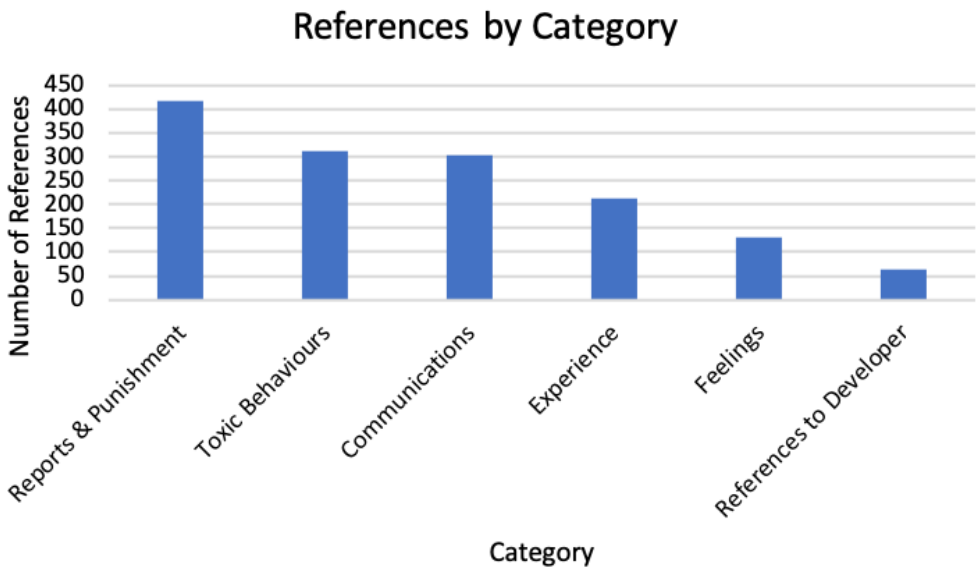

FIGURE 3.2: References by category

- **Verbal abuse** (67) - Acting in an aggressive or hostile manner towards other players in the game via communication methods. This is commonly referred to as 'flaming'.
- **Hate speech** (23) - Using unacceptable language, usually racism or homophobia, via in-game communication methods.
- **Exploits** (11) - Using third party tools or known exploits to alter the game and aid your performance.
- **Team killing** (10) - Deliberately killing other members of your own team. Often referred to as 'tk'.
- **Refusal to communicate with your own team** (14) - Intentionally ignoring methods of communication within the game used by your own team.

These behaviours were defined through a combination of the categories available to report users for in several popular competitive video games (*League of Legends*, *Overwatch*), literature which attempts to categorise behaviours [40, 8] and terminology used within the chosen scraped posts. At this time these are all ranked with equal severity. Analysis of the data gathered should give some insight into the prevalence of each behaviour within communities. Further study

is needed to establish a level of severity from each gathered from community feeling in a more directed manner.a

Terminology

It is worth noting that many of the toxic behaviours identified have colloquialisms and slang associated to them. This demonstrates that these behaviours are commonplace within games to the point of players developing common, simple terms to refer to them. Some terms are colloquialisms, such as 'inting', whereas others are slang, like 'flaming'. These behaviours being so commonplace over many years has allowed for the proliferation of these terms. More official terms which are used by behavioural guidelines are very infrequently used by the community to describe the same offences. Roughly two thirds of toxic behaviour instances use the common term instead of the official behavioural guidelines term. The slang, colloquial terms are clearly more popular in community discussion. This is why terms such as 'trash' were included in the original keywords. Using some of these other colloquial terms was considered when choosing the original key terms. This was not done as many of the terms are not inclusive of all game communities and so could introduce a skew in the final data towards the communities of certain titles.

Behavioural Guidelines

Players seem to find provided behavioural guidelines very hard to understand and apply to their own behaviour. Around half of the relevant posts contain personal stories or chat logs from a previous game played by the writer. These are usually posted to gather community opinion after receiving a punishment. These punishments vary from a short chat ban to a permanent ban on the account. These logs and stories are posted so community input and opinion can be gathered on the behaviour shown. Posters often want to know whether the community find the punishment received to be justified. This seems to show that players find the provided guidelines to be overly simplistic and vague or simply struggle to understand them. They reach out to other members of the community to tell them what is and is not acceptable to breach this uncertainty.

This could be because when players search for guidance from the companies themselves they find it to be vague and unhelpful. One of the posts reviewed included a conversation with a support agent of Riot games who was asked about toxicity. They said this in response:

"Toxicity is a loose term that many players think is just a system that catches on to your swearing, but it's a lot more complicated than that. You see - toxicity is negativity in all it's forms, from the early GG's and sarcastic congratulations, incessant whining in chat, harassment or just comments that distract rather than help, flame rather than aid, or telling people to report someone. Keep in mind that it isn't limited just to the things I mentioned here, but this is to give you a better idea of it. Although I understand where you are coming from and all the frustration involved, please also understand that we have a zero tolerance policy towards toxicity as in there is no excuse for being toxic to a player."

The agent appears to say that toxicity is any form of negative behaviour and that they have zero tolerance. This would mean that any negative behaviours of any kind are not tolerated within matches. It does, however, not help a classify toxic behaviours as players are required to when reporting another player. Nor does it help a player understand what is punishable and what is not. A zero tolerance policy would seem to imply that any negative behaviour is punishable, however, in many other instances Riot have stated behaviours are negative and reportable but not punishable. An example of this is asking for members of your own team to be reported in all chat [31]. Riot have never quantified or expanded upon this to explain to the community why this is the case or how a behaviour may be negative and reportable but not punishable.

Behaviour Monitoring & Reporting Systems

A significant number of questions are also posed regarding the reporting and banning system of games. There are more references to reports, bans and the punishment system than toxic behaviours themselves. Writers ask the community for answers regarding the punishments for various offences. Many seem to feel disheartened and worried because they do not know what to expect as punishment: "This is why we lost the avoid me system, and this is why we are

going to lose our reporting system too if drastic change is not achieved soon this is why we can't have nice things. The only question is how many times I'm going to be unfairly suspended before Blizzard fixes it." If they behave in a toxic manner in a game they are worried what punishment they may receive. They also question if their reports matter when they receive little to no feedback afterwards. Users believe that they get punished when reported whilst others whom they report remain unpunished due to a lack of feedback. There is a strong feeling that users want more transparency and feedback regarding reports and punishments. This appears to be due to a lack of feedback provided when action is taken on a report. Users are also concerned because they feel they are not suitably informed of what is and is not acceptable behaviour: "Blizzard, please explain. Can I play non-meta heroes? Am I toxic? Do I have to join voice chat and listen to these mostly-toxic people? Can I mute them? Can I hide text chat? What is going on?"

Players also complain that sometimes using reports is difficult because of behavioural restrictions. In *League of Legends*, it is a reportable offence to ask for someone to be reported. This is often done to make the enemy team aware of toxic behaviours occurring in the other teams chat. Sometimes examples are given to justify this which can make the person giving the example reportable themselves. Several posters also claim that the report system can be very easily manipulated in this way: "I'm incredibly confused how someone gets permabanned for asking enemy team to report an inter and opening mid, but an intentional feeder doesn't get a ban." There is strong evidence that some punishments are based on a dictionary of unacceptable words. Multiple instances of players manipulating others into using these terms by using variations of known offensive terms. This is done in the hopes of having another player repeat the known term. They can then be reported at the end of the game and likely punished for its use. Users feel that sometimes context is important and needs to be taken into account when punishments are issued, in direct opposition to the previous statement from Riot Games.

In-Game Communication

In-game communication is referenced almost as much as toxic behaviours, with 304 references. It is clear that communities of many games consider refusing to communicate as a toxic behaviour in itself. The in-game text, and sometimes voice depending on game, is where much of the toxic behaviours can be seen. Chat logs clearly show the use of offensive language and harassment of individuals in games. Several words used in these logs would be considered socially unacceptable in an average conversation: "So i got a 14 day ban for saying the word "faggot". It was in no way directed to anyone."

Threats given are often also extremely serious and socially unacceptable such as wishing cancer upon a family member. The facility to mute individual players in chat is provided. This can be used to remove a toxic player from your in-game chat. A punishment employed by some games also mutes a player from speaking at all during matches for a limited time. Muted players, however, sometimes express concern over being reported for refusing to communicate when muted. They feel it is important that a game indicates this mute to other team members to avoid further sanctions.

Feelings & Reactions to Toxic Behaviours

Many posters explicitly express their disappointment in having toxic behaviours present within matches. Multiple posters state that it makes the game less fun for them. It is also stated that toxic behaviour makes users less inclined to return and play the game again. Posters say that having a toxic player, especially someone who intentionally feeds, ruins their game experience. Several posters also comment that toxic behaviours directed towards themselves or their friends greatly diminishes their experience.

Some posters who have been toxic, however, state they have found being toxic to be fun in itself. This is often in response to another player being toxic so they respond in kind. Others have just decided to 'troll' and stop taking the game seriously in response. This behaviour has a fine line with toxic behaviours but are not always reportable offences. Players state they create their

own fun by behaving in this manner. The difference between toxic behaviour and trolling behaviour is very unclear and varies by individual opinion. What is and is not toxic also varies in the same way with some being far more tolerant than others. Trolling behaviour can also be referred to as bad manners, or 'bming'.

3.6.2 Automatic Analysis

An automatic analysis was also conducted on the relevant posts by analysing the use of language with LIWC 2015. Relevant posts were imported into the software and then processed using the LIWC Internal Dictionary 2015 [53]. This dictionary was used as it covers at least 80% of language used from a test data set. The language which is not covered is mostly specific to game mechanics or toxic behaviours which cannot be assigned to pre-existing classifier variables. Some of these also cannot be assigned to categories as where they should belong categorically is part of this research. This generated values for various aspects of the post context, title and post text, such as emotional affects and mental processes. These values can be evaluated to gain insight into the mindset of the poster whilst they were creating their post. This mindset should be reflective of their feelings towards the experience they are detailing or even generally towards the game itself. This data should help to further reinforce any conclusions about a posters feelings determined from the thematic analysis. A greater understanding of the language used, and the poster themselves, provided through psychological analysis of the language may also generate differing conclusions. Summary measures are used to quantify the emotional affect presented in the posts:

The shown measures are defined as follows:

- **Analytic** - A score below 50 indicates a less analytical, more narrative writing style. A score above 50 indicates more analytical writing.
- **Clout** - A score below 50 shows a more humble writing style. A score above 50 shows a writing style which shows high expertise.

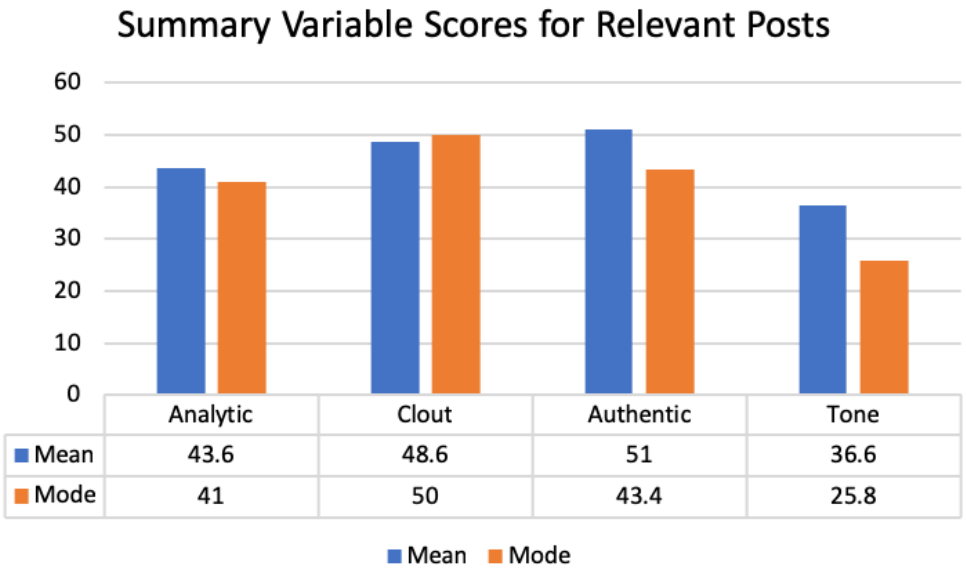

FIGURE 3.3: Summary variable scores for relevant posts

- **Authentic** - A score below 50 indicates guarded, distanced disclosure. A score above 50 indicates honest disclosure.
- **Tone** - A score below 50 shows anxiety and hostility. A score above 50 indicates positivity.

LIWC also returns some other measures which are not directly comparable to these. They may, however, provide further insight on the data in certain scenarios. These are:

-	Words per Sentence	Words > 6 Letters	Dictionary Word Count
Mean	22.9	14.5	83.6
Mode	17.2	12.8	88.2

TABLE 3.1: Question 6 responses

These further measures are defined as follows:

- **Words per sentence** - The mean number of words used in a sentence throughout the post.
- **Words > 6 letters** - The number of words over 6 letters in the post.
- **Dictionary Word Count** - The percentage of words in the post included in the dictionary used for analysis.

Affect Tone

One measure showed that most posts did not show a particularly positive or negative affect. In fact, posts overall displayed slightly more positive affect than negative affect but by a negligible value of 0.3. The level of swearing was also very low, however, this can be attributed to the community guidelines which can restrict the usage of such language. Half of the measured mean negative affect was determined to be anger which demonstrates some posters are clearly angry about aspects of the toxic behaviour they experienced. Posts overall appear to have either been very balanced between positive and negative affect or have shown little to no emotional affect either way. This neutral mean shows that some posts are very positive, others are very Negev but in general posts are usually fairly neutral. It was expected that there would be a measured trend towards negative affect and so this was a surprising result. This may be due to some of the extracted posts being impassioned pleas to the community with very positive overall affect.

It is more likely that the former is true, posters are balanced between positive and negative affect, as the score for the tone is low with a mode of 26. If the posts were lacking of emotion or ambivalent then tone would be very neutral overall for the posts. As the mean tone is low this demonstrates a greater level of sadness, anxiety and hostility within the posts. Hostility was definitely noted in the thematic analysis as many posts are users who were punished for toxic behaviour attempting to defend their position. The mode tone is also noticeably lower then the mean tone, over 10 points, meaning that there are some extremely negative posts tonally. This negative tone likely reflects peoples negative feelings on toxic behaviour and matches in which they are subjected to it. It could also show the negative feelings that people have about games that can result in them displaying toxic behaviours themselves as some posts are from the individual who was toxic during a match. This was an expected result as negative, toxic behaviours have been shown to increase stress levels and decrease in-game performance of those exposed to them. Individuals become

stressed and frustrated when exposed to toxic behaviours an enter a more negative mindset.

Length

The majority of the posts are fairly short with a mean of 327 words and a mode of 123 words. Most longer posts, about one fifth, are so because they consist of a large amount of copied and pasted chat logs. These logs are either recorded by the individual or provided as evidence of punishment when a player has had sanctions applied to them by the game. Long log files explain why the mean word count is approximately 2.7 times the mode. Mean and mode for clout is neutral with the mode being exactly 50. This shows that the posts are not written in a style that makes them sound experienced or humble particularly as there is a balance between them. The balance, again, is likely caused by the experienced players discussing the gameplay whilst also being unsure about received punishments or reports.

Formality Authenticity

The mean analytic score is below baseline with the mode even more so. This score means that the style of writing is more informal, personal and reflects a narrative style of thinking. It is not surprising as posts tend to be made in an informal fashion to be shared with peers of similar interests. The narrative style of thinking makes sense as most of the posts are used to build a narrative, a story, to inform others of events. Narrative style allows users to convey their experiences and feelings in a way others can understand and empathise with. Many posts are created looking for empathy towards the user who has been punished for their deviant behaviours.

Authenticity is, again, mostly around the baseline score and so balanced. The mean slightly favours more authentic language which would have been expected from the honest, personal accounts given. A high authentic score would indicate some form of high exposure and vulnerability from the post author. The mode, in fact, favours a more guarded and distanced style of writing. It

is possible that instead of appearing to be vulnerable as they have been punished, many authors in fact become defensive. This defensiveness means that most posts are slightly on the lower side of the metric. As posters often post to gather the opinions of their peers this is not surprising as they post, knowing they have been punished, and ready for a fight. Some of the more authentic posts are pleas from community members who have experienced but not taken part in toxic behaviours and may have been injured by them.

Summary

Overall the data shows a great deal of variety over the summary variables with some mode values differing by small amounts and others over 10 points. This large variance may be caused by the differing purposes of posts. Some posts are written by players facing sanctions for toxic behaviour whilst others are players pleading to stop toxic behaviours from unreported players. The greatest variances between mean and mode are seen between the values for authentic, difference of 8, and tone, a difference of 11. A large variation in tone was expected as many of the posts are written for different reasons and to express different concerns. Authors often appear to post positively and hopeful when asking the community to behave better or requesting they monitor their own behaviours. Posts dealing with individuals who have received punishment tend to be more negative and downhearted, likely as they have been punished and do not see any way to rectify this. There are some outlier results such as more negative expressions of lost hope in the community from a toxic experience in a recent game or positive users posting about being reformed from past toxic behaviour.

The largest observable variation between mean and mode is for the authentic summary variable. The mean is almost the baseline of 50 but the mode is 11 points more negative. There are more posts which are above baseline but a few posts that are extremely negative which accounts for this. 59 posts are above the baseline of 50, meaning they are more authentic and honest. This leaves 39 posts below the baseline which appear less authentic. The highest scores for authenticity appear to be for pleas and user stories of toxic behaviour, either by the author or another individual. The much more negative posts tend

to include in-game chat logs. As most of these are examples of toxic behaviour where users have been reported and punished for this behaviour, it makes sense for these to be highlighted in this way. These chat logs display some very negative tone, with 4 examples scoring below 2 points for tone showing very few positive qualities to the writing.

These LIWC results show a balanced usage of language in the analysed posts. Some posts were found to be much more hostile and sad than ambivalent which has been shown in previous studies [18]. Some posts were also found to be more informal in style and personal rather than formal [34]. These results were both expected. Most other metrics, however, were found to be very neutral. Posts were not particularly negative or positive, particularly truthful or exposed and demonstrate a very neutral tone. It was expected that posts would be more negative, exposed and show a more negative tone for these factors [15]. This may be due to the wide variety of posts evening out over the small sample size of under 100 posts from the time period. It could also be due to these toxic behaviour accounts not affecting the individual as much as was predicted.

3.6.3 Insights

The results from the thematic analysis proved that discussion of toxic behaviours and the behaviours themselves feature frequently in conversation amongst communities. Community members discuss reporting systems and punishments more than toxic behaviours themselves in any of their forms. This shows they are clearly invested in these systems and aware of their presence in their games, whether reported or reporting another. So much community insight being available may lead to some information discussed amongst communities being useful to make these flawed systems more effective. Many more details about experiences were given when compared to feelings about experiences as posts tried to be objective and state the facts. This nature of writing gives further credence to the idea that these detailed experiences may contain data which is useful to improve existing systems. This large amount of catalogued experiences like these creates a mass of knowledge which gaming companies could utilise to better their own internal systems and improve experiences for all players.

The linguistic analysis proved to lack any solid correlation for the most part, without a strong preference towards either aspect. The variance from the centre point for analytic, clout and authentic is not high enough to carry significance. Tone is an exception to these other variables. Mean for tone is measurably lower than the midpoint and the mode is significantly further so. Despite the posts being fairly neutral for other variables they manage to convey a more negative tone on average. Some posts which are negative tonally also contain posts written in a significantly negative tone as the mode is so low. Community members clearly associate toxic behaviours and behaviour systems as a very negative experience. They can, however, discuss these experiences in an honest manner with narrative writing and moderate levels of expertise.

Community members clearly demonstrate an understanding of toxic behaviours and categorised them clearly within the chosen posts. They reflected on these experiences negatively but with moderate expertise and told them in a narrative fashion. These categories reflect those established by existing game rules so communities appear to have agreed to use these categories themselves. Community members, especially those being punished, appear to often reject rulings based on these previously adopted categories. This rejection may be due to defensive response when guilty, demonstrate a lack of understanding or care for the rules in place. Community members who are not being accused often express opinions either agreeing with existing systems and rulings or arguing against what they see as unfair. These more impartial opinions could definitely be useful to including community opinions in an improvement process for existing systems.

3.6.4 Critical Analysis

It is apparent from the results of the linguistic analysis of the study data that the method was largely inconclusive for gaining insights. This is with the exception of the tonal aspect of the scraped posts which was much more conclusive. There may be several reasons for this.

The dictionary used for post analysis covered around 80% of language used

in the posts. Initially this was deemed sufficient coverage to be able to gain insights. This coverage may not be sufficient to correctly separate posts based on language alone. Many of the most important words in the posts may have been missed. Those terms related to toxic behaviours or colloquial terms commonly used inside these communities may have been missed from the library. This could cause the key, important parts of the posts to not be understood by the analysis and thus be ignored. Modifying a custom dictionary which included terms frequently found within these communities may have helped to return much richer data and lead to more insights.

Many of the posts chosen for analysis were very short. The shortest posts were excluded before being analysed however the threshold for this may have been too low. Longer posts with more language used could have lead to more solid trends within other aspects of the posts. This would, however, have caused an issue with a very low amount of posts being analysed as the particularly long posts were the outliers within the data set. It may be necessary for a more heavily processed and larger data set to be used to find more apparent trends within this kind of post. This data would have more posts and those used would be longer on average to provide more data for analysis. This increases the chance that differences would show between these more verbose posts.

3.6.5 Next Steps

These passive observations of gaming communities have identified many insights into such communities and toxic behaviours. The analyses show that communities at large often discuss such issues. These posts are made by victims, perpetrators and passive observers alike. Additional information could be gathered in a more directed fashion to cover gaps in knowledge seen in this analysis. There is limited categorisation of the various behaviours encountered in the presented posts. Also, there is limited expression of emotions associated with these behaviours from different sides of the incidents. Data on games which are not toxic could also be useful to contrast with games identified here so differences can be compared.

Chapter 4

Study 2

4.1 Introduction

Study one was designed to gather community experiences involving toxic behaviours in a passive manner. This was done in an attempt to gauge community feelings on toxic behaviours. The next step was to design a study which more actively probes community members on the specifics of toxicity to help create a well-rounded understanding. This study is intended to actively ask community members for feedback on toxic behaviours. Data gathered in this fashion can build upon the passive observations made in the previous study and fill in identified gaps in knowledge. An active approach allows these areas to be probed directly. This can be achieved using semi-structured interviews, surveys or other approaches where the thought process of the participant can be directed.

4.2 Design

The aim of this study is to further inform on community feelings on toxic behaviours in online video games. This additional data allows for reinforcing conclusions from the previous study whilst exploring newly discovered areas of intrigue. Data from the previous study can be used to construct examples which can be used in this study to gather feedback on specific discovered behaviours. These behaviours can be ranked to discover which are considered to be most or least toxic by community members. It was decided that, due to the large variety of toxic behaviours identified, that a large number of examples will

be necessary. These examples could be used in a survey to gather community feelings on specific types of behaviour in a directed fashion.

This, however, poses the issue of having a survey that is too complex or too long. If the survey is too long then it is likely fewer responses will be received as people will close it before completion. To alleviate this issue, it was decided that each user will be presented with a random selection of responses from a larger pool. Providing there are enough responses to provide coverage on each example, this approach is more likely to gather a large amount of data. As there are 10 examples for each behaviour this should provide ample coverage for each of the identified categories of toxic behaviour. There are also 10 examples which display no toxic behaviour according to the rules provided by many games. The shorter, simpler survey this creates should be much faster to complete and appear visually far smaller. This will help to elicit more completed surveys after a user opens the survey and looks at it to determine whether or not to complete it.

It was felt that a pairwise comparison was important to include in the survey. This is because it will allow for a ranking to be created of the most to least toxic examples included within the survey. The chosen examples can be referred back to their initial category and be used to create a toxicity index. This index will allow for a method of ranking toxic behaviours based on the level of toxicity perceived by the community from said behaviour. It will also help to justify whether or not this index is accurate using the examples deliberately included which are not inherently toxic. These include being bad at a game or accidentally performing an action to the detriment of your team. Hopefully this index will provide some basis for understanding the severity of certain toxic behaviours according to game communities.

4.3 Methodology

The inductive nature of the first study was focused on observing the community behaviour in regards to toxic behaviour. These observations allowed for patterns to be seen from the observations. Patterns formed distinct categories

of the observed behaviours and general community feelings on toxicity. This second study can now be performed using deductive reasoning based on the previous study. The directed survey question approach allows for further testing on the established categories and community opinion. Statistical analysis of the responses should allow for conclusions to be drawn from community opinion with a high level of certainty. A greater understanding will be provided of community consensus on these toxic behaviours and their prevalence. This should also allow for more aspects of toxic behaviour categories to be understood such as the severity of each behaviour in the communities eyes.

Using posts from study one, a collection of generic examples of toxic behaviours which were observed could be created. A selection of 70 example behaviours based on the previous study was created. This contains 10 examples for each of the six categories of toxic behaviour discovered from the previous study. These include: a player using homophobic slurs; a player killing a team mate deliberately; a player intentionally assisting the enemy team. An additional 10 examples which are not, by game rules, toxic behaviour were also added. These include players perhaps not being good at the game or being a new player who is unsure how to proceed.

The 70 examples can be used in a survey to allow ranking of each from most to least toxic. This pairwise comparison of generic examples of toxic behaviour will allow for an index to be created from the data. An index of this type will identify which types of behaviours community members find to be the most and least toxic. The examples can also be used to determine which behaviours, if any, community members feel demonstrate toxic behaviour. A large sample size from the survey should allow for an overall community consensus to appear on these identified categories of toxic behaviour from the previous study. This survey was shared to many of the subreddits where the data for the previous survey was gathered to further inform previous conclusions or present new insights from the same communities.

4.4 Ethical Considerations

Further work using data from online sources raises more ethical questions to be addressed [20]. All moderators on subreddits used to post this study were asked for permission before any posts were made. This is with the exception of */r/SampleSize* which exists specifically for sharing various studies and questionnaires. This meant that the number of subreddits where the study could be posted was more limited. The study was only posted on */r/SampleSize*, */r/DotA2*, */r/csgo* and */r/Rainbow6* as these are subreddits where express moderator permission was given.

4.5 Implementation

A bespoke web survey was created using HTML5, CSS3, JavaScript and JQuery to facilitate the study. This approach was chosen as each time the survey is loaded, five examples from the pool are chosen and presented to the participant. Presenting only 5 responses per participant ensures that the survey is not long and/or overwhelming to complete. These 5 examples should provide a reasonably varied coverage of the categories from the possible pool of examples. A large collection of responses and the fact there are 10 responses per category should provide enough data to form a meaningful index for toxic behaviours. If over 100 responses are gathered, it should be able to filter these categories down further still into subcategories.

Other questions were also added to the survey which are more generic but fill in gaps left by the coverage of the previous study. Users are asked to categorise their generic examples into the categories derived from the previous study. Before this, they are asked to describe the examples in plain text based on their understanding of the example. This should help to reinforce conclusions made on what categories exist and what they entail. Users are also asked to identify which gaming communities they have personally experienced to be the most and least toxic. These are free form answers from the participant and

not limited to the communities where the study was shared. This should allow for further understanding of the prevalence of toxic behaviours based on a games genre or other factors.

After the survey has been completed, local browser storage is used to prevent a user from taking the survey multiple times. The data gathered by the survey is stored in a Firebase database. This JSON database is separated into each individual response which is assigned a unique identifier. Each response to the given questions is stored within this answer as sub-nodes. This allows for each response to be scrutinised seperately and links to be made between an individuals answers. This approach to storage also allows for responses to be analysed automatically if required as the format is standardised.

The survey was hosted using GitHub pages and accessible from a standard, short web address. This was done as GitHub was used to allow for easy version control during the development of the site. After it was ready to be shared, it was simple to make the page live as the survey was already stored on GitHub. The link was then easily shared on the Reddit sites given above. This survey was also shared on another Reddit page specifically for academic studies and with friends and family. The wide coverage ensures that enough responses were gathered to form meaningful conclusions. Using the randomised example approach necessitates a reasonable level of responses to form informed conclusions with enough data.

4.6 Survey Structure

Each time the survey is loaded, five random examples are selected from the seventy examples available. Any example based questions present the participant with these five random examples. The five examples are the same throughout all questions of a generated survey for one participant. All questions, with all options present, are presented to all participants.

1. Would you consider the following example toxic behaviour?

Example:

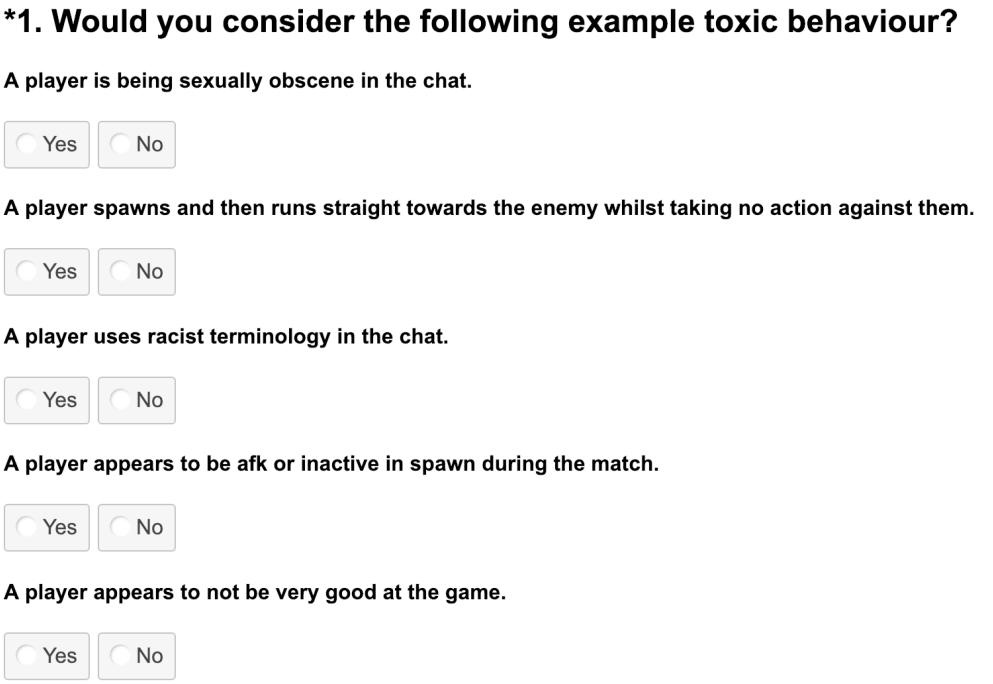

FIGURE 4.1: Question 1 example

The participant is presented with five randomly selected examples of behaviour. They are asked to select either 'Yes' or 'No' for each given example. This allows each participant to specify whether they personally feel each example demonstrates toxic behaviour or not.

2. Order the examples from most toxic to least toxic by dragging them into order.

Example:

The participant is presented with their randomly selected examples in a random order. They are asked to order the given examples from the most to least toxic in their personal opinion. This is done by dragging the given examples into the desired order with most toxic at the top of the items.

3. If you had to describe the type of toxic behaviour (if any) shown in the given example, how would you describe it?

Example:

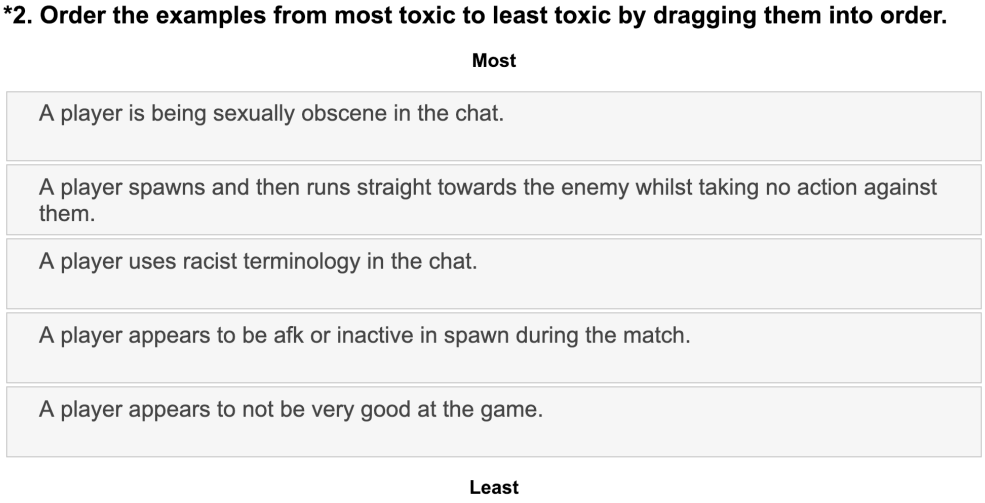

FIGURE 4.2: Question 2 example

The participant is presented with a text box for each randomly selected example they are provided. They may use this text box to provide a free-form description of the example behaviour from their perspective.

4. Here is a list of categories occasionally used to describe toxic behaviours by the community. Order these from most toxic to least toxic.

Example:

The participant is then asked to rank established categories of toxic behaviour in a more structured fashion. They are presented with a list of these established categories in a preset order. The participant may drag the examples into their desired order based on their personal opinion. The top example is the most toxic, descending down the list to least toxic.

5. If you had to categorise the toxic behaviour shown in the examples using the community descriptions, which would you choose? Choose all that apply.

Example:

The participant is then asked to categorise their randomly selected examples in a more structured fashion. They are presented with a selection box for each of the established categories of toxic behaviour from the previous question. The participant may select any number of these categories based on how they would

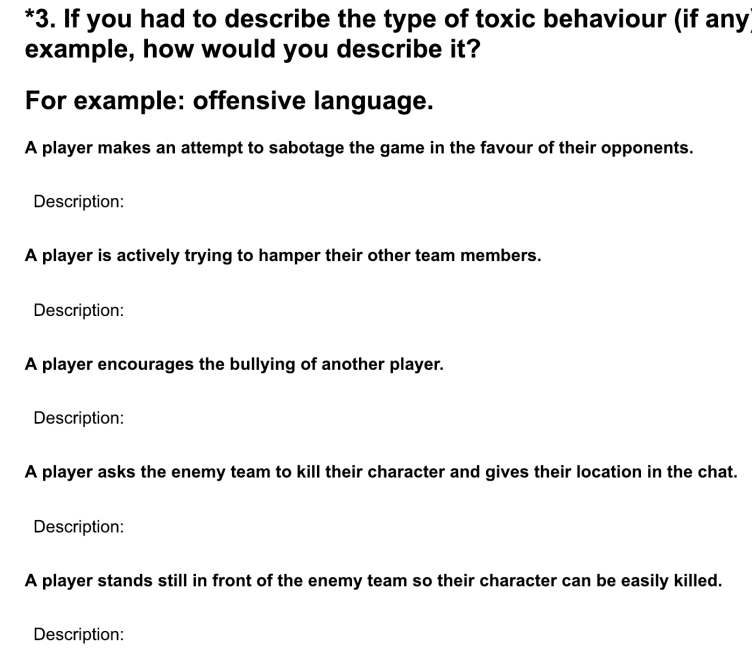

FIGURE 4.3: Question 3 example

classify the example personally. They may also select none if they do not believe any of these categories apply.

6. Which toxic behaviours, if any, from the community definitions have you personally experienced? Choose all that apply.

Example:

The participant is presented with the previously established categories of toxic behaviour. They may select any number of these categories to indicate if they have personally experienced any of them. They may also select none if they have not personally experienced any of these behaviours.

7. Which gaming community have you found to be the most and least toxic?

Example:

The participant is presented with two free-form text boxes. They are asked to name the most toxic and least toxic gaming community they have personally experienced. They may type the name of any video game into each box.

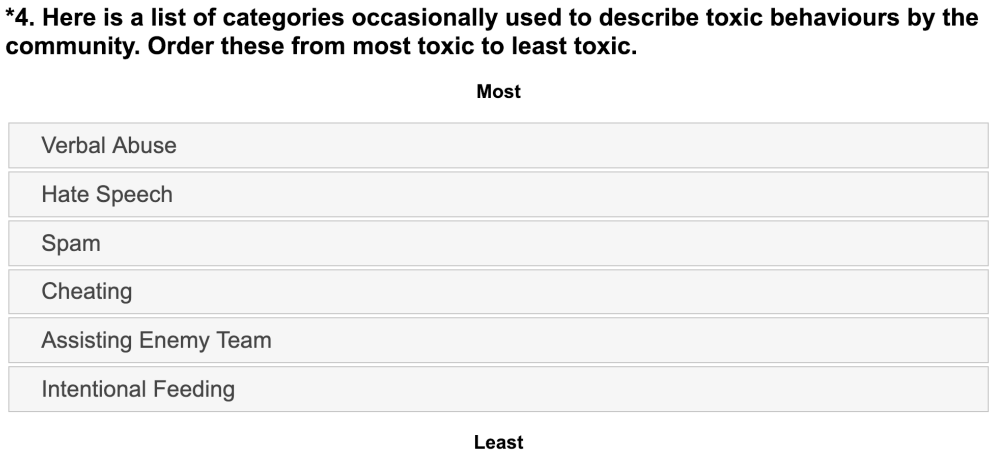

FIGURE 4.4: Question 4 example

4.7 Results

In total, this survey received 165 unique responses over the period of a month. The survey was presented as a web form built using JQuery UI components. Each participant is allowed to complete the survey once and local storage is used to prevent repeated entries. The 5 examples given to each participant are randomly selected from 70 examples inspired by study 1, as intended. Some questions are related to participants general feelings on toxic behaviours by category. Other questions allow participants to give free-form answers to detail their personal experiences. The first section of questions are mostly based on the examples given to the participant. Later questions are focused on toxic behaviours in general.

This survey was posted on several Reddit pages related to video games. These are mostly the subreddit pages used to gather data for study one as long as the page moderators granted permission. These locations are most likely to have users who have personal experience with toxic behaviours. Past experience of toxic behaviours are necessary for completing the survey as it relies on personal experience for later questions. The survey was also shared on several Reddit pages specifically for sharing academic studies. It is hoped that some users here may be more eager to answer a survey and some may have experience with toxic behaviours. This was done to ensure as many valid responses to the survey as possible to built the data set. The survey was also shared within

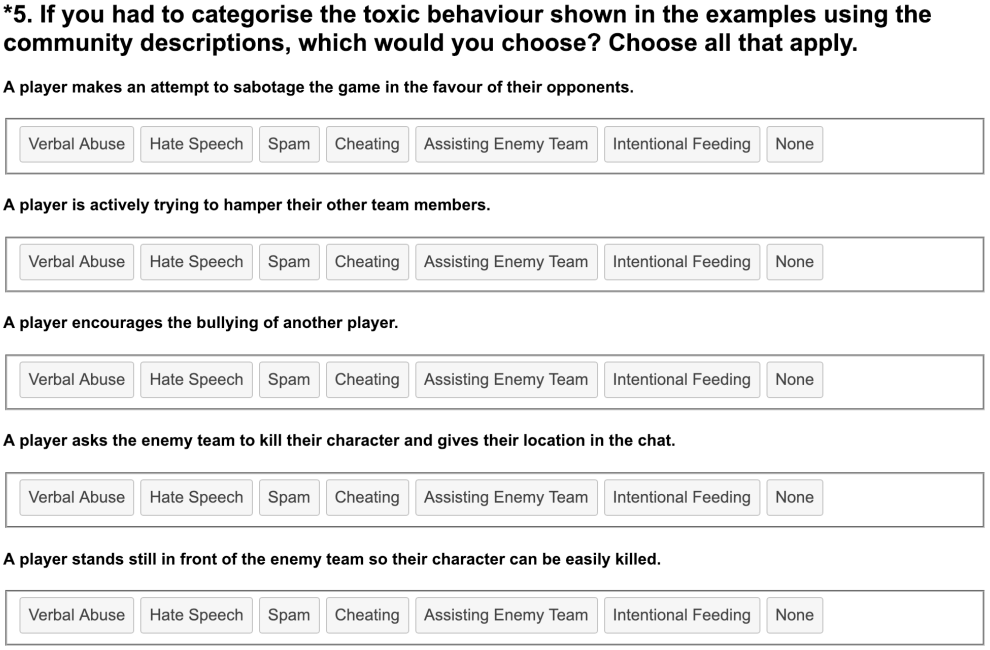

FIGURE 4.5: Question 5 example

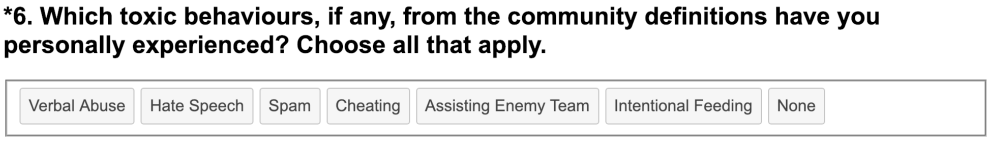

FIGURE 4.6: Question 6 example

gaming communities I am a part of which contain family and friends. It was only shared in groups I knew were actively video game players who participate in matches of one or more of the identified games from study one.

Responses to certain questions provided statistical insight into the mindset of the participants relating to toxic behaviours in general. Other questions help to confirm or deny assumptions made from study one data used to create this survey. Users also made suggestions in free from answers which may provide novel approaches or examples not previously considered. Statistical outliers and correlations were identified.

4.7.1 Question One

The first question asked participants to state whether a given example was toxic or not. The results of this question are as follows:

FIGURE 4.7: Question 7 example

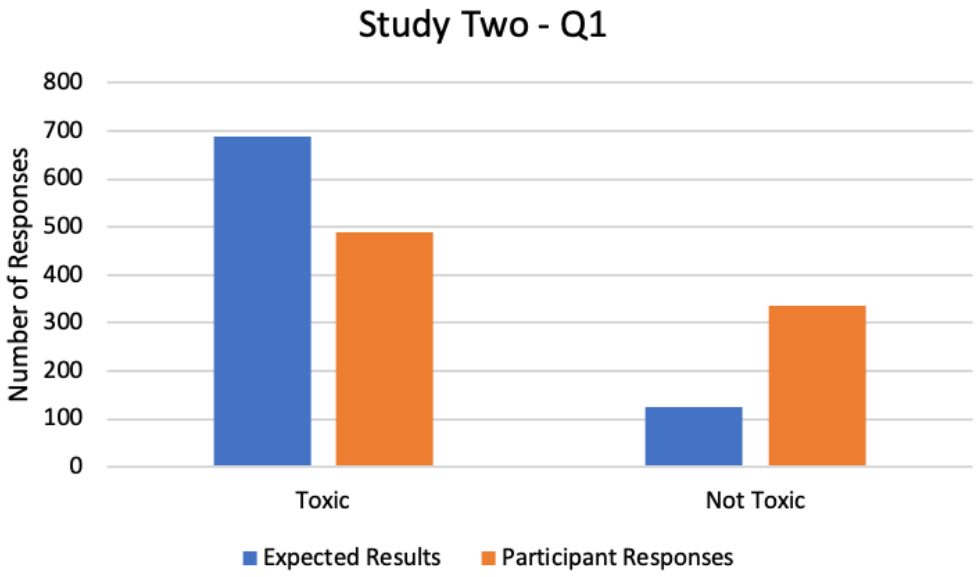

FIGURE 4.8: Question 1 responses

Throughout the 165 surveys generated, 126 examples were presented which were taken from the no toxic behaviour category. In the survey results, however, community participants voted that 336 examples were not displaying toxic behaviours. This means that community members stated that they did not find 210 toxic examples to display toxic behaviours. As these examples were crafted to display toxic behaviours gathered from study one observations these findings are surprising. Study one indicated that the community found these experiences to display toxic behaviours which contradicts several study two responses. The toxic behaviours primarily identified as non-toxic were verbal abuse and hate speech. It was noted that these behaviours are both not directly impactful to gameplay within a match and hosted entirely within the games communication methods.

This could be for several reasons. The examples constructed for study two were crafted to be very generic and with little context to them. This is to allow for the behaviours to be the focus of the behaviour and to allow them to be

understood without background knowledge. Community guidelines presented for video games also often state that context is irrelevant and the behaviour is all that matters. Community members have already shown to be more lenient than game rules provided and not always agree with them. These findings could show that this is not the case for the community. Community members may rate toxic behaviours differently depending on the context of the match in progress and the behaviour itself.

This irregularity could also indicate that community members are simply inconsistent in their views. Survey participants in the second study may not have the same views as those who posts were gathered in study one. Communities have been shown to be very divided on their opinions on toxic behaviour and so this is a possibility. It is also possible that the examples are flawed by design and so community members did not understand them correctly. As they were informed from real examples discovered in study one this seems somewhat unlikely. The process of making the examples generic, however, could have caused the point of the examples to be lost.

Some examples generated to be toxic, however, were rated with a significant favour given to not being toxic. Multiple examples of this can be seen in examples provided for cheating. When compared to answers to question 3 for these examples it is clear that participants did identify the example as cheating but still selected not toxic. This may be due to the difference of opinion demonstrated in study 1 and in literature on cheating as a toxic behaviour. Some examples of spam also heavily favour not toxic for this question which may indicate players find this behaviour annoying but not inherently toxic.

4.7.2 Question Two

Question 2 showed a large variety in the most highly rated examples. As each participant was presented with five examples chosen at random they were only able to provide ratings for the small selection they were presented with. Assigning scores to each example based on its placement in the list and then accounting for how many times the example appeared allows for comparison with the wider results for this question.

Of the top ten examples ordered by their score for this question, the categories were as follows:

- Three examples of verbal abuse
- Three examples of assisting enemy team
- Two examples of hate speech
- One example of intentional feeding
- One example of spam

The first examples which were created to not display toxic behaviour are placed at positions 56, 57 and 58 out of 70. In question 1, all of these were rated to be 100% not toxic. Due to the random nature of the example generation, this appears to be due to the toxic examples not being generated enough in the presented surveys to reach significance from the number of responses. This does, however, demonstrate that the responses provided to be toxic appear higher overall in the order than those which were not. This demonstrates that the toxic examples presented do fit the intended purpose and indicate a toxic behaviour of some description to most participants. There is always some personal influence on these however as some people, as shown in study 1, do not believe certain categories tested are inherently toxic.

Examples of hate speech appear no lower than position 33 in the final order, with half of the examples being in the top 20. Other toxic categories have more of a range and so appear to be less consistent throughout the list:

This consistency within the list is representative of the consistency in community feelings towards the toxic behaviour categories. Verbal abuse, for example, has a range of 46 between the first and last entry in the list and so is far less consistent as a category. The examples lacking toxic behaviour are the most consistent with a range of 13 so participants consistently felt these examples were not toxic. This is good as the examples were chosen to be similar to non-toxic behaviours which are frequently misinterpreted as toxic in study 1. The average player may be more diligent and understanding of toxic behaviours than was previously shown and able to determine this more clearly.

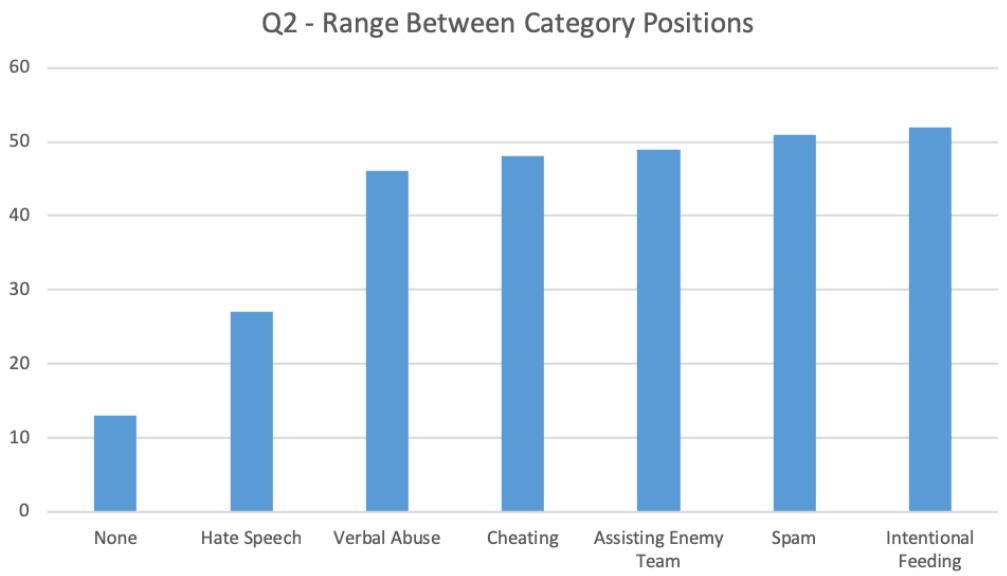

FIGURE 4.9: Question 2 responses

4.7.3 Question Three

For question 3 it appears that participants frequently commented the category name they felt the example belonged in before being asked by the survey to categorise them. This demonstrates that players are often already familiar with the categories used which are defined in many existing games. Some other terms somewhat synonymous with toxic appeared frequently in comments such as 'griefing'.

Several comments question whether or not an example is toxic as they feel that the intent of the is relevant. An example of this: "I'd say this only applies if the person appears to be playing badly deliberately, especially if they're high level" Some examples may be genuine mistakes or unintentional which the player views as not toxic but the same behaviour with deliberate intent would be toxic. This contradicts with the usual way punishments are enforced in games and general rule sets which consistently say that context is irrelevant. Players seem to often disagree with this and place importance on the context of the displayed behaviour.

A few responses on hate speech and verbal abuse indicated that they felt these behaviours were worse as they have real world repercussions. This is likely players acknowledging the difference between personal toxic behaviours

and gameplay related toxic behaviours. Participants often termed personal toxic behaviours as harassment. In the case of hate speech, comments often identified the type of hate speech on display such as racism or sexism.

Some comments relate displayed toxic behaviours to feelings of players at the time they were behaving in such a manner. A participant commented "usually that's just frustration which is more sad than anything" on an example of verbal abuse. They believe this toxic behaviour is being directly caused by the player in question being frustrated at the time and needing to express this feeling.

Another participant said on an example of verbal abuse "Normal in most competitive games". This demonstrates just how normal these behaviours are in online competitive video games [41]. Direct verbal abuse at another individual is viewed as normal behaviour in one of these games by this participant. This is drastically different to most other interactions with other people where this behaviour would be seen as very unacceptable. Toxic behaviours being so normalised may contribute to the prevalence of these behaviours because players feel it is an acceptable way to behave.

4.7.4 Question Four

Question four asked participants to rate the predefined toxic behaviours from study one and established game rules. These toxic behaviour categories were ranked from most to least toxic. The results are as follows:

Of the available behaviours, cheating was most selected as the most toxic option from the provided options. Cheating was selected as the most toxic behaviour by 63 of the 165 participants. The second most selected behaviour for most toxic was Hate Speech with 55 selections. Spam was selected as the least toxic behaviour by 91 participants.

The distribution graph shows a clear preference towards cheating as the most toxic behaviour from responses and spam as the least toxic behaviour. Other categories however display less clear-cut trends to their distributions. Hate speech is fairly evenly distributed between peoples first and second place selections primarily. Verbal abuse is predominantly seen as participants second

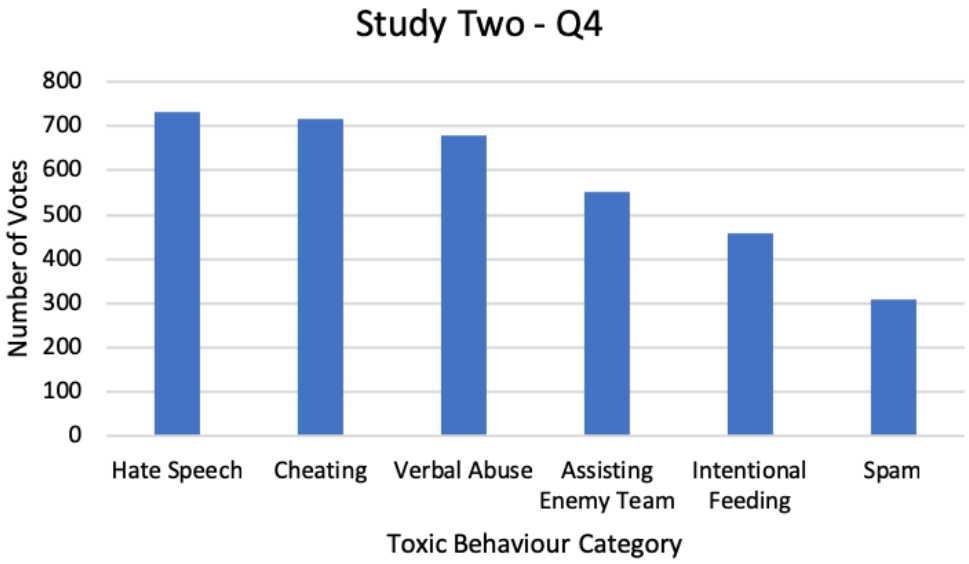

FIGURE 4.10: Question 4 responses

choice despite being third overall. Cheating sees a large resurgence of selections for third most toxic behaviour despite having a very small distribution of peoples selection for second most toxic.

It was expected that spam was the least toxic behaviour overall from the selections available. Spam is widely considered to be less toxic by community members in the previous study with many questioning whether or not it is a toxic behaviour. Several video games have rules which allow for users to be reported for spam alongside other toxic behaviours. This lead to its inclusion in these categories which were inspired by the rules of existing online competitive video games.

Cheating was not expected to be the most selected behaviour after the results of study one. This is because in many genres of online competitive video game using exploits is not feasible. Games in the MOBA genre, for example, are considered almost impossible to exploit due to the complexity of the games mechanics. Past exploits have automated small elements of the game but these are usually very quickly patched by the developer.

This data would imply that even if it is not frequently encountered players consider cheating to be most severe. It is possible that the frequency of toxic behaviour occurring is not relevant to the severity for participants. Cheating is

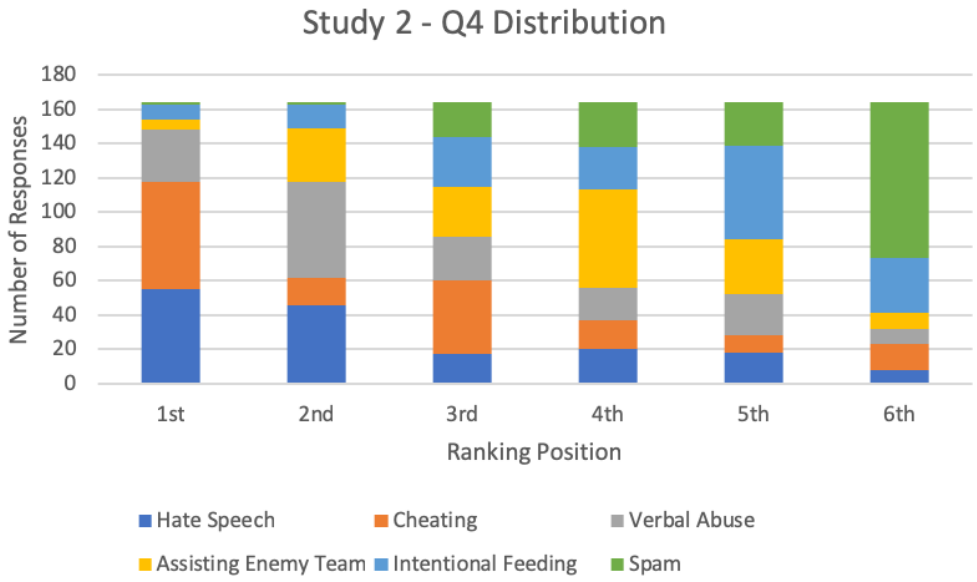

FIGURE 4.11: Question 4 response distribution

an outlier in this way as most other behaviours are very frequently encountered by players given data from the first study. This means it is not possible to draw insights as to the true reason for this. The most likely, however, seems to be that frequency and severity of toxic behaviours are not as linked as was previously assumed for players.

It is possible that community members, regardless of the frequency they experience exploits, still consider them to be the most toxic. Study one raised questions over whether the use of exploits is a toxic behaviour or its own category of behaviour. Again, as the systems to report toxic behaviour also provide the ability to report players for using exploits it was included here. Exploits usually do not allow for any competition within a match as the cheating player has an unfair advantage, It is possible the removal of the competition, a key element of these games, does more harm to a players experience than other behaviours.

Hate Speech amassed more votes than any other category in total from the survey responses. This means Hate Speech is the most toxic behaviour category identified by this study. Verbal Abuse was the third most toxic behaviour

overall. This means that the more personal toxic behaviours, those which affect individuals within the game were thought of as more toxic overall. Personal toxic behaviours amassed 40% more votes than gameplay related toxic behaviours. These gameplay related toxic behaviours, such as Assisting Enemy Team & Intentional Feeding, were overall both less selected. This could indicate that players find personal attacks to be more toxic than gameplay related toxicity.

Answers to this question appear to show correlation between responses and thinking either personal or gameplay related toxic behaviours are more serious. Of the 55 participants who selected Hate Speech as the most toxic behaviour category, 42 of these selected Verbal Abuse as the second most toxic. These participants ranked gameplay related toxic behaviours to be less toxic than these personal behaviours. Other participants rank these gameplay related behaviours to be more toxic than personal behaviours. More participants seem to feel that personal experiences are more toxic and form the majority of responses. The minority find that gameplay related behaviours are more toxic.

4.7.5 Question Five

Question five answers relate to the categorisation of examples by participants. Here is the proportion of examples whose most commonly selected category was the one the example was constructed to fit:

Most categories were largely accurate showing that the constructed examples appeared to participants to accurately represent the behaviour. This was expected as the examples were constructed using data from study 1. There are some exceptions to this, however. Only two of the ten hate speech examples were primarily categorised as hate speech. The other 8 examples were all most commonly categorised as verbal abuse. This may be because participants view hate speech as a subcategory of verbal abuse. This could also be caused by participants being unsure what hate speech is as no definition was provided. This lack of definition is also the case on the behaviour reporting form of most competitive online video games.

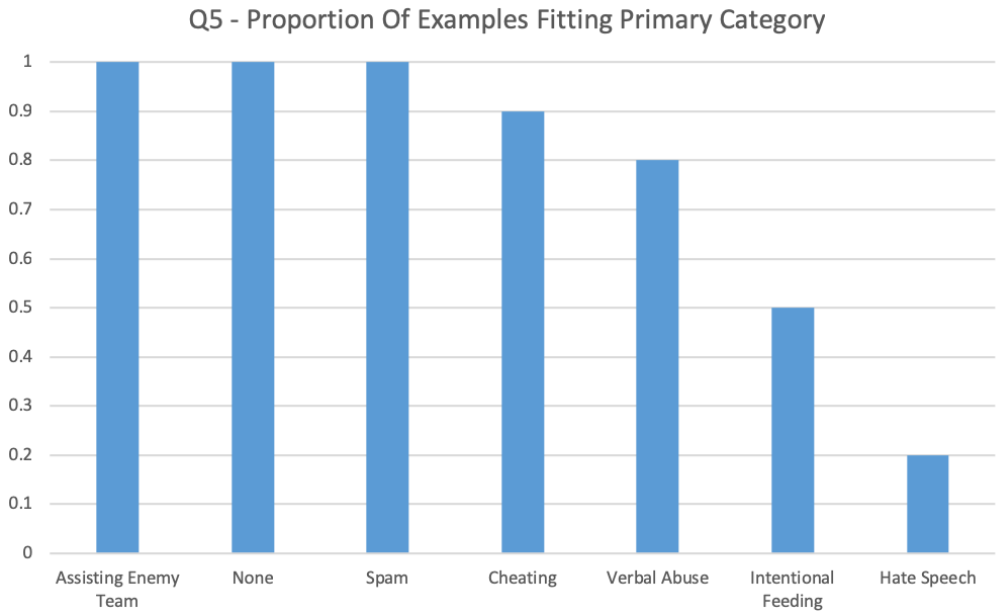

FIGURE 4.12: Question 5 responses

Answers often have the same participant categorising examples as multiple behaviours. Examples which demonstrate hate speech are frequently marked as both hate speech and verbal abuse and vice versa. Examples of intentional feeding are often flagged as both assisting enemy team and intentional feeding. This may be because participants view intentional feeding as one of many ways a player can assist the enemy team. This seems to imply that there is a lot of overlap between the categories that games use to categorise behaviour. Some categories, such as intentional feeding, could be viewed as sub categories of the others. This may be the reason that many games allow you to select multiple categories when sending a behaviour report.

4.7.6 Question Six

Question six asked participants to list the categories of toxic behaviour that they have personally experienced. Responses to this question were as follows:

Of the 165 participants, 68% of participants admitted to experiencing the least reported toxic behaviour which was Spam. 85% of participants admitted to personally experiencing Verbal Abuse in a competitive online video game. Verbal Abuse was the most commonly experienced toxic behaviour category

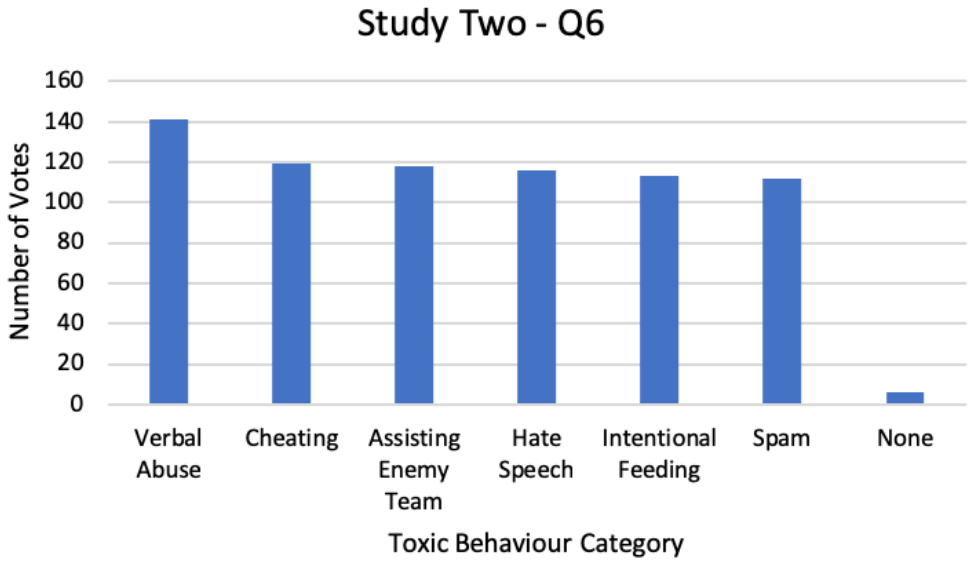

FIGURE 4.13: Question 6 responses

overall. Only 6 of the 165 survey participants stated they had never personally experienced toxic behaviour.

The fact that only 6 of the participants have never experienced toxic behaviours in a game appears to be a very small portion. This, however, may be due to the survey participants being a skewed sample. Few individuals who have not experienced toxic behaviour in a competitive online video game would choose to take a survey on such behaviours. It would also be expected for most players to have experienced some or all of these toxic behaviours. This is because toxic behaviours are so prevalent in these competitive online video games.

Aside from the most and least popular responses, the number of responses for each category are very consistent. There is only a variance of 7 responses between the second and sixth most selected categories of behaviour. This may be because roughly 60% of participants stated that they have experienced every category of toxic behaviour. This demonstrates that all of the identified categories from study one are prevalent in competitive online video game matches. The players active within the community seem to have a well-rounded experience of all previously identified forms of toxic behaviour.

The small variance between all of the central values leads to the significance

of these being very minimal. Statistically these values are not significantly different enough to provide any meaningful insight. All of these values all fall between 69% and 72% of responses experiencing these behaviours. Therefore all of these behaviours can be considered to be experienced in equal amounts. The significant leanings from this question are the prevalence of verbal abuse experienced by survey participants. 85% is an extremely large proportion of responses, even given the expected weighted nature of responses as participants would be expected to have some experience with toxic behaviours. Only 6 responses not having experienced any toxic behaviours is not surprising because of this weighted sample.

4.7.7 Question Seven

Question seven asked participants to name both the most and least toxic gaming communities they have personally experienced. These answers were given as plain text.

For the most toxic community there is a strong trend towards games previously determined as highly toxic. The games used for study 1 data collection appear frequently in user responses. Frequent answers include *League of Legends*, *Counter Strike: Global Offensive*, *Overwatch* and *Dota 2*. Thirty five players identified *League of Legends* as the most toxic community they have experienced. This makes it the most popular choice. Some others users also stated that MOBA titles in general have the most toxic communities. This includes both *League of Legends* and *Dota 2*. Some other games not previously discussed were occasionally identified. These were *Grand Theft Auto Online*, *Dead By Daylight*, *Call of Duty*, *Smite*, *Rust* and *Left 4 Dead 2*

Games identified to have the least toxic communities were frequently not competitive in any way. Users mostly identified games like *Little Big Planet*, *Stardew Valley*, *Animal Crossing*, *Factorio*, *Warframe* and *Mass Effect*. These games are primarily single player or PvE and do not feature direct competition between players as a main element of gameplay. Some games with competitive elements were identified, however. These include games with PvE competitive elements like *Diablo 3*, *Warframe* and *Runescape*. Some games which have PvP

elements but are not PvP focused were also identified. These include *Elite: Dangerous*, *Elder Scrolls: Online* and *Guild Wars 2*. These three titles are more MMO style games with smaller communities and niche appeal. This may lead to more tightly knit, supportive communities with less internal conflict and thus have less toxic interactions.

4.8 Insights

The data from this study provides many insights into the psyche of competitive online video game players and their opinions on toxic behaviour. Feelings on cheating as a toxic behaviour appear to be extremely varied or sometimes even contradictory. In question one, several examples of cheating were chosen to not be displaying toxic behaviour. It can be seen through answers to question 3 that the participant did identify that cheating was taking place in most cases. This would imply that players know what cheating is and can recognise cheating in games however do not feel that it is inherently toxic behaviour. Answers to question 4, however, show more participants identified cheating as their first choice for most toxic behaviour than any other category. Cheating was also identified in total as the second most toxic behaviour from these survey responses. It may also be worth noting that cheating is far less common behaviour than any of the other categories given from data gathered in the first study. However, this frequency may not play a part in participants perception of severity of the behaviour [8].

This anomaly could be interpreted in many different ways. Players may have less experience with cheating in their games as it is less prevalent and so have less of an informed opinion on the category. It is possible that players do not feel that cheating is a toxic behaviour but consider it on another level of bad behaviour. Players may have answered question four by ordering the behaviours they feel are worst to best from the available categories instead of from most to least toxic. Some players may see bad behaviour and toxic behaviour as the same thing in a question like this as all toxic behaviours are by their nature bad behaviours [14]. Another possibility is that as the survey was carried out

by the participant they began to think more about their answers or see toxic behaviour differently. In previous questions, questions were constructed to allow for participants to express their feelings without any other influence. Question 4 may influence participants as it is the first question to provide them with predefined categories to answer with.

Answers to question seven make it clear that there is a difference between players experiences with toxic behaviour in PvP and PvE games. Games often identified as the most toxic are heavily competitive PvP games where players are in direct competition with other players. Competition does seem to be related to the frequency of toxic behaviours [51, 63]. In contrast, games often identified as the least toxic are are often non-competitive PvE games. Games where individuals play alone or without competition are often given as examples [69]. Some games with PvP elements were identified however they are often optional parts of a larger PvE experience within these games. This seems to imply a link between the level of competition between individuals and the level of toxic behaviours present within the game. The more competitive the experience provided by the video game, the more likely a player is to demonstrate toxic behaviour within a given match. Competitive games often also share cooperative elements as you are playing on a team in competition with another team. These cooperative elements may increase toxic behaviour in themselves as a player may disagree with another players choices. Games identified as the least toxic mostly have experiences a player can conduct solo with occasional collaboration and optional competition.

4.9 Critical Analysis

The answers to the survey come from a convenience sample. The survey was shared in the subreddits used in the first study who granted permission for it to be shared. A convenience sample comes with many advantages and disadvantages. Using this sample allowed for the survey data to be collected very quickly and effectively. The participants were ready and waiting to be presented with a survey which is relevant to them. However, using this sample can

also cause issues. There may be bias in the survey responses weighted towards those who play certain games where certain behaviours are more prevalent. If a player has never experienced a game where team killing is possible, for example, it would be very difficult for them to rank against toxic behaviours they have experienced. Players of certain, less popular games are also likely to be underrepresented within the sample.

All participants were required to rank all of the defined categories of toxic behaviour from study 1 in ranking questions. This means that some players who primarily played games where certain behaviours aren't relevant were still required to rank them. These players may have never played a game which featured this type of toxic behaviour if they play few games overall. For players who have experience of all behaviours these questions provide excellent insight into their personal feelings and experiences. However, for players who have not experienced them this question could be weighted towards those they have experienced being rated as worse. They may also have placed these behaviours somewhat randomly in rankings as they do not have a well rounded knowledge of what the behaviour really is. This means the ratings may also be weighted towards participants who played the more popular games when compared to those with less respondents.

Chapter 5

Conclusions

5.1 Results

In both study one and two, it is made very apparently that gaming communities overall share a strong negative opinion of toxic behaviours in their chosen games. Players on forums frequently comment that players who are punished have been given what they deserve with little to no remorse after viewing provided evidence. In the study 2 survey it is apparent that the community feel that toxic behaviours all negatively affect gameplay by comments made. The difference comes from what an individual personally deems toxic behaviour or not toxic behaviour. In general, players seem to be more lenient on toxic behaviours than video game companies as many express interest in knowing the context of the behaviour shown. Without knowing this context participants often chose to label the behaviour as not toxic. Community members therefore seem to place a great deal of value on the context of the displayed behaviour before making a judgement. Existing systems in games ignoring context and games companies often stating that context is irrelevant appears to be in direct conflict with community feelings.

The negative tone that forum users often use when talking about toxic behaviour is made clear by the linguistic analysis in study 1. Most variables provided by the analysis are insignificant as they are very close to the baseline value of 50. Tone is a key exception to this with a mean of 36.6 and a mode of 25.8. Speaking of toxic behaviours is often done in an anxious and hostile way despite being baseline analytical, authentic and clout. Toxic behaviours clearly

elicit anxiety from community members when discussed and the topic clearly makes community members more hostile. This could be one of the reasons that gaming communities have such a negative impression of toxic behaviours aside from experience in game. Even out of game discussions about toxic behaviour bring on anxiety and hostility in community members.

Categories of toxic behaviour defined from literature and study one appeared to be mostly effective for study 2. As expected, the 'Spam' category was widely selected as the least toxic by survey participants with over half of responses indicating it was the least toxic overall. 'Spam' was, however, still primarily deemed to be a toxic behaviour in responses and so still considered toxic by the community. 'Hate Speech' was deemed to be the most toxic behaviour by number of votes however more people chose cheating as their most toxic behaviour in order. 'Hate Speech' was also shown to have a lot of overlap with 'Verbal Abuse' in the eyes of the community with many examples being classed as both. Participants often viewed 'Hate Speech' as a form which 'Verbal Abuse' can take. The 'Assisting Enemy Team' category was much less toxic than predicted before study 2 provided insight into community feelings. Gameplay related toxic behaviours (Assisting Enemy Team, Intentional Feeding) appear to be widely considered in responses to be less toxic than personal toxic behaviours (Hate Speech, Verbal Abuse). This was expected to be more balanced with many participants feeling that these behaviours are more toxic than personal behaviours more frequently. It appears than overall community members are more committed to the well-being of themselves and other community members than the outcome of their competitive video game matches. This was also observed in study 1 with the large amount of discussion and support provided for others on the forums.

Cheating being the most toxic behaviour when ordered is contradictory to example based questions where many identified examples as cheating however selected that this was not toxic. It is possible that cheating being first place as a toxic behaviour is an anomaly due to the style of the question. Question 4 assumes that all given categories are toxic and asks players to rate these categories from most to least toxic. Participants cannot state that a given behaviour

is not toxic in this question and so may be rating the behaviours on their perceived severity as opposed to toxicity Therefore the validity of cheating as a toxic behaviour category appears to be questionable compared to the others. Community members may feel that cheating is a different category of behaviour in itself. It is also worth noting that in study one cheating appears very infrequently compared to other behaviour categories and so communities may be less experienced with it.

Community members themselves seem to fall into two categories - those who feel personal toxic behaviours are more toxic and those who feel gameplay related behaviours are more toxic. Community members who consider personal toxic behaviours to be more toxic are more numerous than those who feel gameplay related behaviours are more toxic. These community members care more about well-being than their teams performance in video game matches and so place well-being above game performance. Competitive video games tend to be regarded as having much more toxic communities so it is possible that competition leads to these more personal behaviours. Players who play more competitive games may care more about their performance in a match than personal behaviours and, so, commit personal toxic behaviours they consider less severe. A wider appreciation of the severity of the personal toxic behaviours seems to be present along with an understanding of the impact these behaviours can have on an individual. This may be because most players have experienced some form of verbal abuse aimed at themselves in a video game or otherwise and so can empathise with victims.

5.2 Research Questions

1. What would be an accurate definition of toxic behaviour, if one is possible?

Toxic behaviour has been shown to be very diverse, complex and open to interpretation throughout both studies. As such, a simple definition is likely not suitable to describe the phenomenon. This may explain why past definitions are

nebulous and do not accurately describe toxic behaviour. Some definitions from past studies are as follows:

- Griefers are characterised as gamers who stalk, hurl insults, extort, form gangs, kill and loot. Overall, griefers can be summed as those who like to cause other gamers to enjoy their game less [13].
- Cyberbullying, griefing, mischief, and cheating often grouped as toxic behaviour however the definition of toxic behaviour is often unclear due to differences in expected behaviour, customs, rules, and ethics across games [41].
- A player who derives his/her enjoyment not from playing the game, but from performing actions that detract from the enjoyment of the game by other players [48].
- Negative in tone, hurtful in intent, mean, profane, and/or insulting. [66].
- Intentional harassment of other players with the intent. Utilising aspects of the game structure or physics in unintended ways to cause distress for other players [70].
- Dependant upon the nature of the game eg. some have killing as not allowed where others have it as the main objective [14].
- General short term aggressive behaviour brought on by the competitive nature of the game [1]
- Behaviours which are socially unacceptable and disrupt other players experiences. The ambiguity of rules motivates players to discuss and evolve their understanding of behaviour. The official forums are a key venue in which players interpret rules [39].
- Deliberately bad in-game performance, offensive language, verbal abuse [8].
- The use of profane language by one player to insult or humiliate a different player in his own team [45].

The number of different categories that toxic behaviour can fit is large and many share some overlap. Some individuals class behaviours like cheating via exploits as a toxic behaviour but other individuals do not. An attempt at a definition may satisfy some but not others due to these varying interpretations of what toxic behaviour is. A key finding of analysis here is two distinct categories of toxic behaviour appear to have distinguished from one another. These are:

- Gameplay related toxic behaviour - behaviours which have a direct affect on a teams performance in the game. These may cause players to directly lose a match from deliberate poor performance or harming team cohesion.
- Personal toxic behaviour - behaviours which are attacks towards an individual or a group rather than gameplay. May indirectly affect game performance. This includes verbal abuse and hate speech towards other players in a given match.

Using these as a basis a more generic definition of toxic behaviour can be derived:

"Behaviours which are outside of socially acceptable behaviours within society. These behaviours either hamper a given match of a video game directly or indirectly. They may also involve personal attacks against individuals or groups."

2. How does a games community feel about toxic behaviour?

Both studies make abundantly clear that gaming communities are unified in a strong disdain of toxic behaviours. Those who are victims of toxic behaviour often post in forums to complain about their experiences and state that it has ruined their enjoyment. Perpetrators of toxic behaviours who have been punished try and plead that they have not shown toxic behaviours but do not argue that toxic behaviours are acceptable. When presented with examples of toxic behaviour, participants in the survey expressed their disapproval of the behaviours in free form comments. Toxic behaviour is clearly a problem that is on the forefront of gaming communities and that they wish as a collective to be dealt with in a more successful manner.

3. Are there different levels of severity in toxic behaviour?

Through the second study, participants were asked to grade examples of toxic behaviour and categories as a whole by level of severity. Some examples and categories were shown to be more toxic than others according to the community responses to these questions. Personal toxic behaviours were shown to be widely considered as more toxic than gameplay related behaviours. Community members felt these personal attacks were more toxic than other behaviours which affected the performance and outcome of their video game matches. Scores for different behaviours were similar but some are clearly distinguished from others. 'Spam' was clearly shown to be far less toxic than the other behaviours in community opinion yet still considered a toxic behaviour. This indicates that to the community there are different levels of severity in different toxic behaviours.

5.3 Objectives

- Understand the systems currently in place to deter toxicity in competitive online video games.

Existing systems in place to monitor behaviour come in two variants - entirely automated or community reliant. Entirely automated systems often use simple dictionary analysis to pick up on banned terms and automatically issue punishments to offending players. They may also monitor statistics about gameplay which may be suspicious and use this with player reports to issue punishments. The nature of these systems means context is irrelevant as the system cannot understand any context given in communication or in statistics by how the match was conducted.

Community reliant systems work through community members deemed as respectable by a set of criteria being given the opportunity to view game information and vote on a players conduct. These systems are usually majority rule with the suspect being either punished or pardoned based on the outcome of

this voting. These systems allow community members to have say in what is acceptable behaviour in their chosen game and these individuals can understand context. The participating community members, however, are often provided with strict guidelines to adhere to in their voting which limits their ability to influence a games rules.

- Understand previous studies into the effects of playing competitive online video games.

Past studies have attempted to further understanding of toxic behaviours using several varied approaches. Some have attempted to assess the connection between competition versus collaboration and their effects on toxic behaviours demonstrated. Others have tried to relate the prevalence of toxic behaviours to other external factors such as gender. More generalised studies have been conducted into the psychology of bullying and other abusive behaviours that are akin to personal toxic behaviours. Few studies have attempted to gather information on toxic behaviour from an entirely community driven perspective which could provide further insights into toxic behaviour as a whole.

- Find or derive definitions of toxic behaviour from a wide array of sources.

Both past studies and video game developers have given definitions for toxic behaviour in the past. These definitions are usually nebulous and do not provide much insight. Definitions provided by game developers to guide community members were observed to be so vague that any negative behaviour of any kind regardless of context is classed as toxic behaviour. Past academic studies have attempted to define toxic behaviour using a simple definition however these do not usually adequately convey the nuances of the concept. They often miss vital elements such as classification, severity or importance of varying personal interpretations of toxic behaviours.

- Discover what form/s toxic behaviour can take.

Study one combined with past literature helped to clearly identify different forms that toxic behaviour can take. These were split into personal toxic behaviours and gameplay related toxic behaviours. Toxic behaviours were then

split into categories using existing online competitive video game rules and toxic experiences spoken of in forum posts. Some of these categories feature varying degrees of overlap, appearing to almost be subcategories of other behaviours with more specific requirements.

- Discover whether there are different severities of toxic behaviour.

Conclusions from studies show gaming communities believe that toxic behaviours come in varying levels of severity. Some behaviours are viewed as more trivial whereas others are viewed as very detrimental to either gameplay or any victims of more personal behaviours. These more trivial behaviours are, however, still viewed as toxic behaviour by most participants in study two.

5.4 Grand Challenges

- Large Playerbases

This challenge is unlikely to go away at any time soon and so will need to be dealt with as appropriately as possible. The size of existing playerbases has contributed to the prevalence of automated punishment systems in existing online competitive video games. Games with millions of active players make manual review systems for behaviour unrealistic to implement. This problem is only likely to grow as the playerbases of video games grow and new video games are released which attract new players. Since large playerbases cannot be changed, any future solutions to the problem of toxic behaviour must accommodate for them.

- Inconsistent Community-based Systems

Existing community driven systems are used to allow for a more personal touch to behaviour management whilst allowing for community members to help dictate what behaviours are not acceptable. Community members may also understand the context present in the data they are given in a way that automated systems cannot. This context is viewed as important by the community for accurate judgements, as shown in study two. These systems are, however,

prone to differing interpretations by the community members who cast votes using these systems. These differing interpretations were made clear in study two as different participants expressed different feelings towards the same categories of behaviour. Community-based systems may need to establish further guidance for participating players as to what the expectations are for behaviour in the game. This could be accomplished using a test of given examples much like the survey in study 2 with correct and incorrect answers. These systems could also be adapted to take note of majority community feeling and adjust game rules according to consensus. This would allow committed players to influence the future of the game and players behaviours.

- Inaccurate Automated Systems

It is clear from study results that community members value context and feel this is vital for accurate judgements on given examples of behaviour. Existing automated systems cannot understand context and entirely ignore it. This allows for the systems to enforce game rules which usually specify that the context of the behaviour is irrelevant. This directly contradicts with established community feelings on context and punishment. To fit with what the community appear to want, rules and therefore these automated systems would need to be revised. Rules would need to accept that context is relevant in the case of toxic behaviour and allow for some judgement to take place on a case by case basis. The automated systems which enforce these rules would then need to be modified in some way to account for this rule change. Depending on the level of consideration allowed this could involve a significant commitment to development of more intelligent systems. Other solutions such as manual reviews of player reports are infeasible due to other grand challenges.

5.5 Future Work

Several opportunities for future work have been demonstrated throughout the conducted studies. A clear difference between PvP and PvE video game titles was demonstrated in the final question of study 2. Several of the PvE titles

named as least toxic by participants feature optional PvP elements. A study could be conducted into the behaviour of players who do and do not participate in these optional PvP experiences. A comparison between the two groups of players within the same video game could help to further establish the effects of competition on toxic behaviour.

A more in-depth study focused on comparing and contrasting players opinions on personal toxic behaviours versus gameplay related toxic behaviours could prove fruitful. This study would allow for a clear line to be drawn between the two and community opinion on each to be assessed seperately. A clearer definition between the two could help to further solidify understanding of the different forms toxic behaviour can take. These more overarching categories could also help to alleviate the identified issues with overlap between existing categories if further understood and used.

Several pieces of literature assessed the possible impacts of competitiveness on aggression and other factors [51, 46, 1]. There is room for further work directly investigating this competitiveness and toxic behaviours. This could help to establish if there is a more direct relationship between competition and toxic behaviour. This competition could have a direct impact on the category, frequency or severity of any toxic behaviours displayed.

Some of the analysis performed here has helped to develop further understanding of gaming communities feelings on toxic behaviour as a whole. There is further room to expand on this with a study aimed specifically at players who do demonstrate toxic behaviours identified here. These toxic players may have very different feelings towards the behaviours as perpetrators of them. Further understanding of these individuals could help to further the understanding and prevention of toxic behaviours as a whole.

Some games are shown to display different types of toxic behaviours and to different frequencies depending upon changes in gameplay mechanics. There have been past studies which have attempted to identify differences between genre and player enjoyment [47]. There is space to perform such an analysis on the differences between toxic behaviours and different genres of video game. This could help improve understanding on a genre by genre basis to help focus

future efforts to understanding and controlling the rampant toxic behaviours present in many of the games featured.

Appendix A

Study 2

A.1 Study 2 Examples

Verbal Abuse

- A player verbally abuses another player in the chat.
- A player tells another player to physically harm themselves in the chat.
- A player encourages the bullying of another player.
- A player asks for another member of their team to be reported for toxic behaviour.
- A player harasses a member of their team in the chat.
- A player threatens to physically harm another player in the chat.
- A player threatens to harm another players family members in the chat.
- A player joins in with bullying another player in the chat.
- A player advocates being unlawful in the chat.
- A player harasses a member of the enemy team in the chat.

Hate Speech

- A player questions another players intelligence using a derogatory term in the chat.
- A player uses homophobic terminology in the chat.

- A player uses racist terminology in the chat.
- A player uses sexist terminology in the chat.
- A player is sexually objectifying another player in the chat.
- A player is being sexually obscene in the chat.
- A player is being discriminatory to another player in the chat.
- A player uses transphobic terminology in the chat.
- A player is told they should play a specific role because of their gender.
- A player is told they should play a specific character because of their gender.

Spam

- A player repeatedly sends the same message in the chat.
- A player on your team is repeatedly pinging the same objective.
- A player is advertising third party services in the chat.
- A player is advertising their stream in the chat.
- A player is repeatedly interrupting the chat with irrelevant information.
- A player repeatedly disrupts communications in the chat.
- A player will not allow anyone else to engage in the chat.
- A player is intentionally speaking over others in the chat.
- A player provides links to a third party site in the chat.
- A player interrupts conversations repeatedly with advertisement.

Cheating

- A player on your team appears to know your location through obstructions.
- A player appears to be dodging everything without a single failure.

- A player has not missed a shot or ability all game, even when great effort is made to evade.
- A player appears to be hitting 100% headshots for the entire match.
- A player appears to be exploiting a bug to their own benefit or the benefit of their team.
- A player appears to be using a third party tool to automate a facet of gameplay.
- A player is actively engaging in an activity to provide themselves an unfair advantage.
- A player appears to not be taking any damage despite being hit with damaging shots or abilities.
- A player appears to be teleporting around the map in an impossible fashion.
- A player appears to be dealing impossible amounts of damage in a very short time.

Assisting Enemy Team

- A player is telling their opponents your location in all chat.
- A player appears to be collaborating with the enemy team.
- A player keeps killing other members of their own team.
- A player is actively trying to hamper their other team members.
- A player appears to be AFK or inactive in spawn during the match.
- A player intentionally uses their characters abilities to cause your in-game death.
- A player intentionally helps the enemy team to gain an objective.
- A player is wasting all of their character abilities without being close to any opponents.

- A player is intentionally telling the enemy team when you are attempting to take an objective.
- A player makes an attempt to sabotage the game in the favour of their opponents.

Intentional Feeding

- A player is repeatedly committing in-game suicide with their character.
- A player appears to be letting the enemy team kill their character.
- A player states in the chat they have given up and run at the enemy to kill their character.
- A player requests your team surrenders then begins to die intentionally.
- A player stands still in front of the enemy team so their character can be easily killed.
- A player goes AFK in an exposed location where their character can be easily killed.
- A player spawns and then runs straight towards the enemy whilst taking no action against them.
- A player asks the enemy team to kill their character and gives their location in the chat.
- A player dies in-game and then gives up dies intentionally for the rest of the match.
- A player decides to let their character die because they believe that their team is not helping them.

None

- A player misses a key ability in a fight.
- A player appears to not be very good at the game.
- A player failed to dodge an enemy shot or ability.

- A player asks for someone to help them in the chat.
- A player makes a suggestion of a different strategy or approach.
- A player unintentionally allows the enemy team to take an objective.
- A player accidentally causes your death.
- A player requests you to choose a different character.
- A player is playing a character who does not fit the current metagame.
- A player attempts to coordinate your team's actions in the chat.

A.2 Study 2 Questionnaire

*1. Would you consider the following example toxic behaviour?

A player is being sexually obscene in the chat.

Yes No

A player spawns and then runs straight towards the enemy whilst taking no action against them.

Yes No

A player uses racist terminology in the chat.

Yes No

A player appears to be afk or inactive in spawn during the match.

Yes No

A player appears to not be very good at the game.

Yes No

*2. Order the examples from most toxic to least toxic by dragging them into order.

Most

A player is being sexually obscene in the chat.

A player spawns and then runs straight towards the enemy whilst taking no action against them.

A player uses racist terminology in the chat.

A player appears to be afk or inactive in spawn during the match.

A player appears to not be very good at the game.

Least

***3. If you had to describe the type of toxic behaviour (if any) shown in the given example, how would you describe it?**

For example: offensive language.

A player makes an attempt to sabotage the game in the favour of their opponents.

Description:

A player is actively trying to hamper their other team members.

Description:

A player encourages the bullying of another player.

Description:

A player asks the enemy team to kill their character and gives their location in the chat.

Description:

A player stands still in front of the enemy team so their character can be easily killed.

Description:

***4. Here is a list of categories occasionally used to describe toxic behaviours by the community. Order these from most toxic to least toxic.**

Most

- Verbal Abuse
- Hate Speech
- Spam
- Cheating
- Assisting Enemy Team
- Intentional Feeding

Least

***5. If you had to categorise the toxic behaviour shown in the examples using the community descriptions, which would you choose? Choose all that apply.**

A player makes an attempt to sabotage the game in the favour of their opponents.

Verbal Abuse | Hate Speech | Spam | Cheating | Assisting Enemy Team | Intentional Feeding | None

A player is actively trying to hamper their other team members.

Verbal Abuse | Hate Speech | Spam | Cheating | Assisting Enemy Team | Intentional Feeding | None

A player encourages the bullying of another player.

Verbal Abuse | Hate Speech | Spam | Cheating | Assisting Enemy Team | Intentional Feeding | None

A player asks the enemy team to kill their character and gives their location in the chat.

Verbal Abuse | Hate Speech | Spam | Cheating | Assisting Enemy Team | Intentional Feeding | None

A player stands still in front of the enemy team so their character can be easily killed.

Verbal Abuse | Hate Speech | Spam | Cheating | Assisting Enemy Team | Intentional Feeding | None

***6. Which toxic behaviours, if any, from the community definitions have you personally experienced? Choose all that apply.**

Verbal Abuse | Hate Speech | Spam | Cheating | Assisting Enemy Team | Intentional Feeding | None

***7. Which gaming community have you found to be the most and least toxic?**

Most toxic:

Least toxic:

Bibliography

[1] Paul JC Adachi and Teena Willoughby. "The effect of video game competition and violence on aggressive behavior: Which characteristic has the greatest influence?" In: *Psychology of violence* 1.4 (2011), p. 259.

[2] Fair Play Alliance. *Fair Play Alliance*. 2018. URL: http://www.fairplayalliance.org/.

[3] Craig A Anderson and Brad J Bushman. "Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature". In: *Psychological science* 12.5 (2001), pp. 353–359.

[4] Craig A Anderson and Melissa Morrow. "Competitive aggression without interaction: Effects of competitive versus cooperative instructions on aggressive behavior in video games". In: *Personality and Social Psychology Bulletin* 21.10 (1995), pp. 1020–1030.

[5] Craig A Anderson et al. "Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review." In: *Psychological Bulletin* 136.2 (2010), pp. 151–173.

[6] British Esports Association. *Education and esports: Which colleges and universities are offering courses?* May 20, 2019. URL: https://britishesports.org/news/education-and-esports-which-colleges-and-universities-are-offering-courses/.

[7] Brock Bastian, Jolanda Jetten, and Helena RM Radke. "Cyber-dehumanization: Violent video game play diminishes our humanity". In: *Journal of Experimental Social Psychology* 48.2 (2012), pp. 486–491.

[8] Jeremy Blackburn and Haewoon Kwak. "Stfu noob!: predicting crowdsourced decisions on toxic behavior in online games". In: *Proceedings of the 23rd international conference on World wide web*. ACM. 2014, pp. 877–888.

[9] Gary D Bond and Adrienne Y Lee. "Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language". In: *Applied Cognitive Psychology* 19.3 (2005), pp. 313–329.

[10] Virginia Braun and Victoria Clarke. "Using thematic analysis in psychology". In: *Qualitative research in psychology* 3.2 (2006), pp. 77–101.

[11] Geir Scott Brunborg, Rune Aune Mentzoni, and Lars Roar Frøyland. "Is video gaming, or video game addiction, associated with depression, academic achievement, heavy episodic drinking, or conduct problems?" In: *Journal of behavioral addictions* 3.1 (2014), pp. 27–32.

[12] Peggy Chekroun and Markus Brauer. "The bystander effect and social control behavior: The effect of the presence of others on people's reactions to norm violations". In: *European Journal of Social Psychology* 32.6 (2002), pp. 853–867.

[13] Vivian Hsueh-Hua Chen, Henry Been-Lirn Duh, and Chiew Woon Ng. "Players who play to make others cry: The influence of anonymity and immersion". In: *Proceedings of the international conference on advances in computer enterntainment technology*. ACM. 2009, pp. 341–344.

[14] Thomas Chesney et al. "Griefing in virtual worlds: causes, casualties and coping strategies". In: *Information Systems Journal* 19.6 (2009), pp. 525–548.

[15] Kimberly M Christopherson. "The positive and negative implications of anonymity in Internet social interactions:"On the Internet, Nobody Knows You're a Dog"". In: *Computers in Human Behavior* 23.6 (2007), pp. 3038–3056.

[16] Valve Corporation. *Steam: Game and Player Statistics*. Aug. 21, 2018. URL: `https://store.steampowered.com/stats/Steam-Game-and-Player-Statistics`.

[17] Tech Crunch. *Ubisoft now auto-bans Rainbow Six Siege players who use toxic language*. Ed. by Jordan Crook. July 17, 2018. URL: https://techcrunch.com/2018/07/17/ubisoft-now-auto-bans-rainbow-six-siege-players-who-use-toxic-language/.

[18] Matthew S Eastin. "The influence of competitive and cooperative group game play on state hostility". In: *Human Communication Research* 33.4 (2007), pp. 450–466.

[19] Blizzard Entertainment. *Blizzard's Code of Conduct*. Ed. by Blizzard Entertainment. Nov. 25, 2017.

[20] Charles Ess and the AoIR ethics working committee. *Ethical decision-making and Internet research: Recommendations from the aoir ethics working committee*. Nov. 27, 2002.

[21] Eurogamer. *Blizzard's Jeff Kaplan says toxic behaviour is slowing Overwatch development*. Ed. by Vic Hood. Sept. 14, 2017. URL: https://www.eurogamer.net/articles/2017-09-14-blizzards-jeff-kaplan-says-toxic-behaviour-is-slowing-overwatch-development.

[22] Christopher J Ferguson and John Kilburn. "The public health risks of media violence: A meta-analytic review". In: *The Journal of pediatrics* 154.5 (2009), pp. 759–763.

[23] Christopher J Ferguson and Stephanie M Rueda. "Examining the validity of the modified Taylor competitive reaction time test of aggression". In: *Journal of Experimental Criminology* 5.2 (2009), p. 121.

[24] Chek Yang Foo and Elina MI Koivisto. "Defining grief play in MMORPGs: player and developer perceptions". In: *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*. ACM. 2004, pp. 245–250.

[25] PC Gamer. *Former Dallas Fuel Overwatch pro xQc suspended again, might miss World Cup*. Ed. by Samuel Horti. Aug. 12, 2018. URL: https://www.pcgamer.com/uk/former-dallas-fuel-overwatch-pro-xqc-suspended-again-might-miss-world-cup/.

[26] Riot Games. *Will Tribunal Return?* Ed. by Bioluminescense Scathlocke Draggles Fearless. Aug. 23, 2018. URL: https://nexus.leagueoflegends.com/en-gb/2018/08/will-tribunal-return/.

[27] Douglas A Gentile et al. "The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies". In: *Personality and Social Psychology Bulletin* 35.6 (2009), pp. 752–763.

[28] Douglas A Gentile et al. "The effects of violent video game habits on adolescent hostility, aggressive behaviors, and school performance". In: *Journal of adolescence* 27.1 (2004), pp. 5–22.

[29] Peter Holtz and Markus Appel. "Internet use and video gaming predict problem behavior in early adolescence". In: *Journal of adolescence* 34.1 (2011), pp. 49–58.

[30] Riot Games Inc. *The Summoner's Code*. Ed. by Riot Games Inc. Oct. 16, 2017. URL: http://na.leagueoflegends.com/en/featured/summoners-code.

[31] Riot Games Inc. *We don't ban people for asking for reports*. Ed. by Riot NaKyle. Mar. 28, 2017.

[32] Statista Inc. *Reddit users: unique montly visits 2017*. Ed. by Statista Inc. Oct. 27, 2017. URL: https://www.statista.com/statistics/443332/reddit-monthly-visitors/.

[33] Its$_E$*pi*. *Chat Toxicity Update*. Ed. by Ubisoft. Mar. 2, 2018. URL: https://www.reddit.com/r/Rainbow6/comments/81izuv/chat_toxicity_update/.

[34] Adrienne Holz Ivory and Christine E Kaestle. "The effects of profanity in violent video games on players' hostile expectations, aggressive thoughts and feelings, and other responses". In: *Journal of Broadcasting & Electronic Media* 57.2 (2013), pp. 224–241.

[35] Daniel Johnson and John Gardner. "Personality, motivation and video games". In: *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. ACM. 2010, pp. 276–279.

[36] Daniel Johnson, Lennart E Nacke, and Peta Wyeth. "All about that base: differing player experiences in video game genres and the unique case of moba games". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 2265–2274.

[37] Barbara K Kaye and Barry S Sapolsky. "Taboo or not taboo? That is the question: Offensive language on prime-time broadcast and cable programming". In: *Journal of broadcasting & Electronic media* 53.1 (2009), pp. 22–37.

[38] Ben Kirman, Conor Lineham, and Shaun Lawson. "Exploring mischief and mayhem in social computing or: how we learned to stop worrying and love the trolls". In: *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2012, pp. 121–130.

[39] Yubo Kou and Bonnie A Nardi. "Governance in League of Legends: A hybrid system." In: *FDG*. 2014.

[40] Haewoon Kwak and Jeremy Blackburn. "Linguistic analysis of toxic behavior in an online video game". In: *arXiv preprint arXiv:1410.5185* (2014).

[41] Haewoon Kwak, Jeremy Blackburn, and Seungyeop Han. "Exploring cyberbullying and other toxic behavior in team competition online games". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 3739–3748.

[42] Joel D Lieberman et al. "A hot new way to measure aggression: Hot sauce allocation". In: *Aggressive Behavior* 25.5 (1999), pp. 331–348.

[43] Jeffrey Lin. "The science behind shaping player behavior in online games". In: *annual Game Developers Conference*. 2013.

[44] Alexandru Losup et al. "Analyzing implicit social networks in multiplayer online games". In: *IEEE Internet Computing* 18.3 (2014), pp. 36–44.

[45] Marcus Märtens et al. “Toxicity detection in multiplayer online games”. In: *Network and Systems Support for Games (NetGames), 2015 International Workshop on*. IEEE. 2015, pp. 1–6.

[46] Winter Mason and Aaron Clauset. “Friends ftw! friendship and competition in halo: Reach”. In: *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM. 2013, pp. 375–386.

[47] Elisa D Mekler et al. “A systematic review of quantitative studies on the enjoyment of digital entertainment games”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM. 2014, pp. 927–936.

[48] Jessica Mulligan and Bridgette Patrovsky. *Developing online games: An insider's guide*. New Riders, 2003.

[49] Joaquim AM Neto, Kazuki M Yokoyama, and Karin Becker. “Studying toxic behavior influence and player chat in an online video game”. In: *Proceedings of the International Conference on Web Intelligence*. ACM. 2017, pp. 26–33.

[50] Newzoo. *Most Played Core PC Games | US & EU*. Mar. 29, 2018. URL: `https://newzoo.com/insights/rankings/top-20-core-pc-games/`.

[51] C Thi Nguyen and José Pablo Zagal. “Good Violence, Bad Violence: The Ethics of Competition in Multiplayer Games.” In: *DiGRA/FDG*. 2016. URL: `http://www.digra.org/wp-content/uploads/digital-library/paper_39.pdf`.

[52] Mai-Hanh Nguyen. *See how much the top eSports teams, athletes, and their organizations make*. Ed. by Business Insider UK. Jan. 11, 2018. URL: `http://uk.businessinsider.com/top-esports-teams-players-salaries-2018-1`.

[53] James W Pennebaker et al. *The development and psychometric properties of LIWC2015*. Tech. rep. University of Texas, Austin, 2015.

[54] Nataliia Pobiedina et al. "On successful team formation: Statistical analysis of a multiplayer online game". In: *Business Informatics (CBI), 2013 IEEE 15th Conference on*. IEEE. 2013, pp. 55–62.

[55] Hanneke Polman, Bram Orobio De Castro, and Marcel AG van Aken. "Experimental study of the differential effects of playing versus watching violent video games on children's aggressive behavior". In: *Aggressive behavior* 34.3 (2008), pp. 256–264.

[56] Polygon. *Former pro Overwatch player xQc banned for 15 days over in-game toxicity*. Ed. by Julia Alexander. Aug. 14, 2018. URL: https://www.polygon.com/2018/8/14/17688098/xqc-overwatch-twitch-suspension-streaming-toxicity.

[57] Rabindra A Ratan et al. "Stand by your man: An examination of gender disparity in League of Legends". In: *Games and Culture* 10.5 (2015), pp. 438–462.

[58] Florian Rehbein et al. "Prevalence and risk factors of video game dependency in adolescence: results of a German nationwide survey". In: *Cyberpsychology, Behavior, and Social Networking* 13.3 (2010), pp. 269–277.

[59] Inc. Riot Games. *IWillDominate Tribunal Permaban & eSports Competition Ruling*. Ed. by Inc. Riot Games. Dec. 4, 2012. URL: http://forums.na.leagueoflegends.com/board/showthread.php?t=2864421.

[60] Inc. Riot Games. *League Players Reach New Heights In 2014*. Jan. 28, 2014. URL: https://web.archive.org/web/20171022073409/https://www.riotgames.com/articles/20140711/1322/league-players-reach-new-heights-2014.

[61] Inc. Riot Games. *The indefinite banning of Tyler1*. Ed. by Inc. Riot Games. Apr. 30, 2016. URL: https://boards.na.leagueoflegends.com/en/c/player-behavior-moderation/8a75KeUR-the-indefinite-banning-of-tyler1.

[62] Victoria L Rubin and Sarah C Camm. "Deception in video games: examining varieties of griefing". In: *Online Information Review* 37.3 (2013), pp. 369–387.

[63] Mike Schmierbach. ""Killing spree": Exploring the connection between competitive game play and aggressive cognition". In: *Communication Research* 37.2 (2010), pp. 256–274.

[64] Jae Youn Shim, Tae Hyun Kim, and Seong Whan Kim. "Decision support of bad player identification in moba games using pagerank based evidence accumulation and normal distribution based confidence interval". In: *International Journal of Multimedia & Ubiquitous Engineering* 9.8 (2014), pp. 13–16.

[65] Kenneth B Shores et al. "The identification of deviance and its impact on retention in a multiplayer game". In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM. 2014, pp. 1356–1365.

[66] Sara Sood, Judd Antin, and Elizabeth Churchill. "Profanity use in online communities". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2012, pp. 1481–1490.

[67] AOL (UK). *'Overwatch' player toxicity is delaying game updates*. Ed. by Matt Brian. Sept. 14, 2017. URL: https://www.engadget.com/2017/09/14/blizzard-overwatch-toxicity-developer-update/.

[68] Ars Technica UK. *Riot rolls out automated, instant bans for League of Legends trolls*. Ed. by Kyle Orland. May 26, 2015. URL: https://arstechnica.co.uk/gaming/2015/05/riot-rolls-out-automated-instant-bans-for-league-of-legends-trolls/.

[69] Kellie Vella, Daniel Johnson, and Leanne Hides. "Playing alone, playing with others: Differences in player experience and indicators of wellbeing". In: *Proceedings of the 2015 annual symposium on computer-human interaction in play*. ACM. 2015, pp. 3–12.

[70] Dorothy E Warner and Mike Raiter. "Social context in massively-multiplayer online games (MMOGs): Ethical questions in shared space". In: *International Review of Information Ethics* 4.7 (2005), pp. 46–52.

[71] Dmitri Williams. "Bridging the methodological divide in game research". In: *Simulation & Gaming* 36.4 (2005), pp. 447–463.

[72] Russell B Williams and Caryl A Clippinger. "Aggression, competition and computer games: computer and human opponents". In: *Computers in human behavior* 18.5 (2002), pp. 495–506.

[73] Michele L Ybarra and Kimberly J Mitchell. "Youth engaging in online harassment: Associations with caregiver–child relationships, Internet use, and personal characteristics". In: *Journal of adolescence* 27.3 (2004), pp. 319–336.